Reduced-Parameter Motion Estimation in PET Using Respiratory Motion Models

Balfour, Daniel Robert Malcolm

Awarding institution:
King's College London

King's College London

Faculty of Biomedical Engineering

Division of Imaging Sciences

Reduced-Parameter Motion Estimation in PET using Respiratory Motion Models

Doctoral Thesis

Author:
Daniel R Balfour

Supervisors:
Paul K Marsden
Andrew P King
Andrew J Reader

January 10, 2017

King's College London

Abstract

Positron emission tomography (PET) is a modality with high temporal resolution but long acquisition times. This can result in blurred images due to subject motion. Respiratory motion in particular is an unavoidable source of degradation, which can cause issues with quantification and clinical interpretation. An important characteristic of respiratory motion is its pseudo-cyclic nature, which has previously been exploited to form mathematical models which describe the motion, driven by a small number of parameters. The aim of this project is to use this form of motion modelling to estimate motion using information acquired from both dynamic magnetic resonance (MR) scans and from the acquired PET data itself to correct for the effects of motion. The use of motion models in this way can overcome the high levels of noise which otherwise characterise the estimation problem.

First the feasibility of using motion models is investigated using synthetic data consisting of individual PET gates simulated using real motion information. The PET gates are registered using constraints provided by a motion model derived from MR images. A novelty of this approach is that this is the first time PET data have been used to indirectly drive a parameterised motion model.

The next part of the project attempts to formalise the motion estimation process by incorporating the reduced-parameter motion model into the PET image reconstruction. An analytical gradient for a single motion parameter that drives the model is derived from the same objective function used to estimate the image. This results in significant noise averaging, providing robustness to the high level of noise typically found in PET data acquired over short time frames. This is shown to improve robustness to noise well enough that the number of gates can be increased and overall motion correction performance improved.

Finally, the formulation is extended further to also model photon attenuation effects. This is shown to improve the performance of the algorithm when dealing with synthetic data that includes attenuation.

Acknowledgements

First and foremost, I would like to extend my deepest gratitude to my supervisors for their invaluable support throughout the course of my PhD. To Paul Marsden for providing me the opportunity to become a student at King's and shepherding me through it. To Andy King for always having an open door, and being patient and kind enough to listen to my smallest questions and thoughts. And to Andrew Reader, whose mathematical tutoring and insights helped me overcome many mental blocks. All of you inspire me to be the best academic I can be.

I owe much to my other colleagues at King's, for their help and also for their interesting and exciting discussions. I especially owe a big thank you to Christoph Kolbitsch for his patience and flexibility when it came to acquiring MR data, and to Harry Tsoumpas and Irene Polycarpou for their support when I was learning the ropes with STIR.

I would like to thank my Mum and Dad, who have always supported me and without whom I would not have been lucky enough to get to where I am today. They provided so many opportunities to me and always encouraged me to be myself. A huge thank you to the rest of my family for supporting me. In particular, I owe much to Geraldine and Grahame Drew for putting me up with a bed, a roof, and TLC; and Michael and Alison Bruce for always being there when I need them. And last, but not least, my grandmother Angela Clement, who has always supported me and brought sunshine into my life.

Finally I'd like to thank my friends, old and new, for supporting me through this PhD. In particular, I couldn't have done it without the constant companionship of Mike Vallance and Lauren Melleney, Stuart Russant for mathematical discussions, my understanding housemate Carys Lees, and especially Daniel Edie, for sticking by my side throughout the highs and lows of PhD studentship.

'Scientific progress goes "boink"?'

— Hobbes, from *Calvin and Hobbes*, Bill Watterson 1991

Contents

List of Tables

List of Figures

Acronyms

AF	Attenuation factor
CT	X-ray computed tomography
DC	Direct-correspondence
FBP	Filtered back-projection
FDG	Fluorodeoxyglucose
FOV	Field of view
FWHM	Full-width at half-maximum
IC	Indirect-correspondence
LOR	Line of response
MAP	Maximum *a posteriori*
MBB	Minimal bounding box
MCIR	Motion-compensated image reconstruction
ME-MCIR	Motion-estimating MCIR
ME-RTA	Motion-estimating RTA
MR	Magnetic resonance
NCC	Normalised cross-correlation
PET	Positron emission tomography
PSF	Point-spread function
ROI	Region of interest
RTA	Reconstruct-transform-add
SI	Superior-inferior
SNR	Signal-to-noise ratio
SPECT	Single-photon emission computed tomography
SSD	Sum of squared differences
STIR	Software for Tomographic Image Reconstruction
SUV	Standardised uptake value
UTE	Ultra-short echo time
VOI	Volume of interest

Chapter 1

Introduction

1.1 Motivation and Objectives

Respiratory motion in positron emission tomography (PET) is an unavoidable source of error in lesion quantification and localisation, which directly impacts disease staging, radiotherapy tracking, and other clinical/research PET applications. Methods have previously been introduced which reduce the parameterisation of respiratory motion estimation without significant reduction in performance using, amongst other methods, magnetic resonance (MR) imaging [1]. This class of motion estimation techniques is referred to as parameterised motion modelling.

This thesis presents the work done towards discovering an optimal strategy for respiratory motion correction of PET in a hybrid PET-MR setting by using parametrised motion modelling to facilitate reliable motion estimation from PET data. Initial work involved reconstruction of respiratory-binned ('gated') PET images, simulated from real volunteer MR data using the Software for Tomographic Image Reconstruction (STIR, [2]). The simulations were subsequently used to estimate the respiratory position of each PET image via a constrained registration algorithm. This has been published as a journal paper [3], and forms the first methodology chapter.

The second methodology chapter discusses and assesses transferring the constrained registration procedure introduced in the first methodology chapter into the PET

image reconstruction itself. Since the motion model used is analytical, it was possible to combine it with Poisson log-likelihood (the objective function used for PET image estimation) to produce an analytical update gradient for motion estimation. This was tested using emission-only PET simulated with measured non-rigid motion transformations.

Since real PET data is not emission-only, the motion correction method in the second methodology chapter was too simplistic, but served as a proof of principle. It was found that inclusion of attenuation into the PET simulations negatively affects the ability to estimate motion using the reconstruction-incorporated method. The third and final methodology chapter examines the possibility of incorporating attenuation into the motion estimation procedure. The analytical gradient is re-derived with explicit inclusion of attenuation into the system model. This produces an additional counter term in the analytical gradient, which corrects the error in motion estimation.

In conclusion, this thesis aims to show that the incorporation of a motion model into the PET reconstruction allows drastic parameter reduction (down to one additional parameter to be estimated per PET gate) in a way that still sufficiently estimates full 3D, non-rigid motion for low signal-to-noise (SNR) images, without any use of the MR scanner during PET acquisition. The prospective challenges of applying this to real data in the future are discussed, with suggestions on future directions this research could take.

1.2 Contributions

The original contributions of this thesis to the field of respiratory motion correction in PET-MR can be summarised as follows:

Using PET Data to Drive Indirect-Correspondence Motion Models

A wide range of respiratory motion correction techniques have been introduced in the PET literature. Early attempts to do this generally involved transformation of individual PET images prior to averaging them to obtain a motion-corrected image. This would often involve some method of image registration, although the difficulty of non-rigid motion estimation and noisy nature of the individual PET gates proved to be a limitation to these methods. Alternatively, other modalities could be used to generate motion estimates. The introduction of simultaneous PET-MR allows the use of co-registered MR imaging, which can be used to acquire a subject-specific set of motion estimates. In previous work, it was shown that motion estimates can be drawn together to make a parameterised motion model, which is able to produce full non-rigid motion estimates by only measuring simpler (but correlated) data, such as the relative position of the surface of the chest [4]. It has also been shown that these models can be used to indirectly estimate motion through optimisation of a small number of hidden variables using imaging data. The work in this thesis is the first to use PET data itself as the data to drive this indirect optimisation method. This has the advantage of only requiring the estimation of the few hidden variables, providing robustness against the noisy quality of the PET images. In this sense, using a motion model in this way can be thought of as an MR-constrained registration of two PET images.

Incorporation of Parameterised Motion Models into PET Reconstruction

Whichever way motion estimates are generated for respiratory motion correction, it was found in previous work that motion correction of pre-reconstructed gated images could negatively affect the quality of the final image [5]. Instead, motion estimates were eventually incorporated into the data model used in PET iterative image reconstruction. More recently, there has been an effort to estimate both motion and activity in the reconstruction at the same time, since this feels a natural solution

to two related optimisation problems. However, this can often involve estimation of many extra parameters in addition to the intensities in the PET image: in the extreme case, J voxel intensities are required plus $3JG$ displacements (one for each of G gates in each direction), although this is usually mitigated using reduced-sampling techniques such as B-spline grids. This thesis proposes a Poisson likelihood–based framework that allows one to replace the many additional motion parameters with a parameterised motion model, forming a strong constraint on the subset of motion transformations that can be estimated. As above, this allows a significant reduction in the number of parameters that need to be estimated to produce full motion estimates for on-the-fly motion correction during reconstruction. Only 1 additional parameter is required per gate, such that, for G gates, $J + G$ parameters must be estimated in total for both image and motion estimation using this technique. This reduction in estimation parameters lessens the technique's sensitivity to noise without the need for regularisation, possibly enabling a higher number of gates in the reconstruction. This can improve the level of motion correction that can be achieved due to less motion within each gate.

1.3 Thesis Outline

This thesis is organised into 9 chapters. Chapter 2 provides the background of PET in general, with focus on reconstruction methods and data correction. Chapters 3 and 4 progress onto the problem of motion in PET, with a literature review of techniques that have been used in the past to estimate and correct motion in PET. Due to the lack of ground truth in the motion correction problem, the experiments in this thesis are based on MR-derived PET simulations. Since this is a feature of all the methodology used here, the main concepts have been collected into chapter 5. Chapters 6, 7, and 8 describe the novel methods in this thesis. Chapter 9 provides a summary of the contributions and describes some limitations of the proposed methods, with an overview of future directions. Finally, conclusions are drawn.

Chapter 2

An Introduction to Positron Emission Tomography

Positron emission tomography is a medical imaging modality that allows full 3D, non-invasive imaging of a subject. This is made possible by the existence of positron-emitting radionuclides. These are administered to patients in tiny quantities in the form of radiotracers: molecules which are labelled with the radionuclide through, for example, substitution of a functional group. The most common example in clinical PET is the substitution of a hydroxyl group (–OH) in glucose with an atom of fluorine-18. The glucose analogue will, up to a point, function the same way metabolically as the original molecule, but it can be detected using the radioactivity of the fluorine.

Positron emission tomography can detect the spatial distribution of these radiotracers at remarkably low concentrations (a typical amount of administered activity is around 400-800 MBq, corresponding to a picomolar concentration of radiotracer). This, when coupled with models of the tracer's role in metabolism, can produce functional information without strong perturbation of the metabolic pathway under investigation.

For functional imaging to be useful, it is important to achieve quantitative accuracy and reliability. Since invention, achieving this accuracy in PET – either through building an image from the data ('reconstruction') or as a post-process – has been

a significant focus of research in the medical imaging community.

In this chapter, the nature of PET will be outlined, focussing on an overview of data correction techniques and a detailed description of reconstruction procedures.

2.1 Basic Principles of PET

2.1.1 Underlying Physics

As mentioned at the start of this chapter, PET is a medical imaging modality which uses radionuclide-labelled metabolites to gain functional information about the subject. The archetypal example is fluorodeoxyglucose (FDG), a metabolic analogue to glucose, which accumulates in respiring cells at a rate proportional to the rate of respiration [6]. The attached radionuclide, fluorine-18, decays in a manner that allows for localisation and quantification of the decay by the PET scanner.

The radionuclides used in PET imaging are exclusively those which emit β^+, or positron, radiation. β^+–decay occurs in nuclides with an excess of protons. It results in the conversion of a proton into a neutron via the weak interaction, emitting a positron and an electron anti-neutrino in the process. This is illustrated in figure 2.1.

The emitted positron continues into surrounding matter at high velocity. It is deflected through elastic and inelastic collisions with local atoms and nuclei, losing energy and scattering as it moves away from the mother nucleus. In the case of ^{18}F, this corresponds to a mean range of 0.22 mm [7], although this depends on the density of the surrounding material. Once enough energy has been lost, it interacts with a nearby electron to form a metastable electron-positron state known as positronium, which eventually annihilates itself [6]. The annihilation of the electron-positron pair produces a minimum of two photons, as dictated by momentum conservation. In the majority of cases where the spins of the particles are antiparallel ('parapositronium'), two annihilation photons are created each with energy very close to that of the rest mass of the electron, 511 keV. Conservation of momentum in the

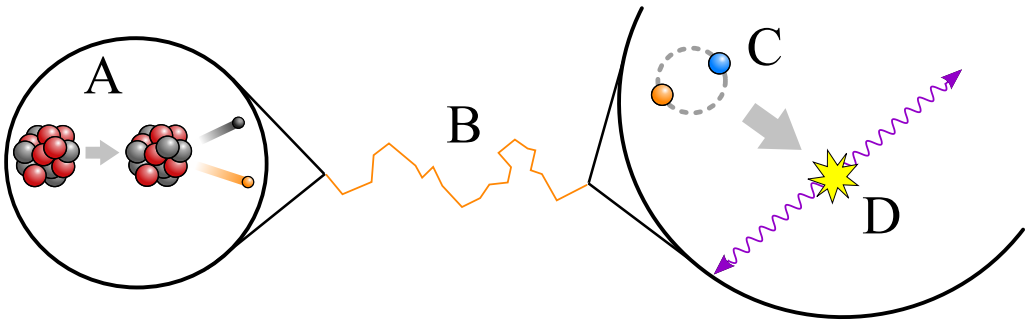

Figure 2.1: **Beta decay and positron annihilation** Positron annihilation is the key underlying process in PET. This process occurs in several stages. First, a nucleus decays by converting a proton to a neutron, producing a positron and a neutrino (A). This positron propagates away from its mother nucleus at high speed. Through multiple collisions it loses energy inelastically (B). Once its kinetic energy is low enough, the positron forms a metastable state with a local electron known as positronium (C). The positronium state eventually self-annihilates, producing, on average, two back-to-back annihilation photons at 511 keV (D). These photons are what we aim to detect in PET, allowing localisation of the original emission event.

positronium rest frame ensures that the photons are emitted 180° apart.

2.1.2 Scanner Instrumentation

When positronium self-annihilates, two 511 keV annihilation photons are produced 180° apart. These continue away from the point of origin, and may be detected by the PET scanner. This is illustrated in figure 2.2. The aim is to detect annihilation events within the field of view (FOV), since these will generally have been produced by the tracer within the subject.

The detectors used in conventional PET are scintillator material coupled to photomultiplier tubes, or PMTs [6]. More recently, avalanche photodiodes (APDs) and silicon-photomultipliers (SiPMs) have been used due to their compactness and ability to operate in a strong magnetic field [8]. These detectors are usually arranged in planes or rings around the FOV.

All PET scanners require the detector to be coupled to some sort of scintillator. A scintillator is usually a crystal composed of a material which, upon excitation by an incident high-energy photon, creates secondary photons in the optical/ultraviolet band of the spectrum. For this reason, it is favourable for scintillation materials to

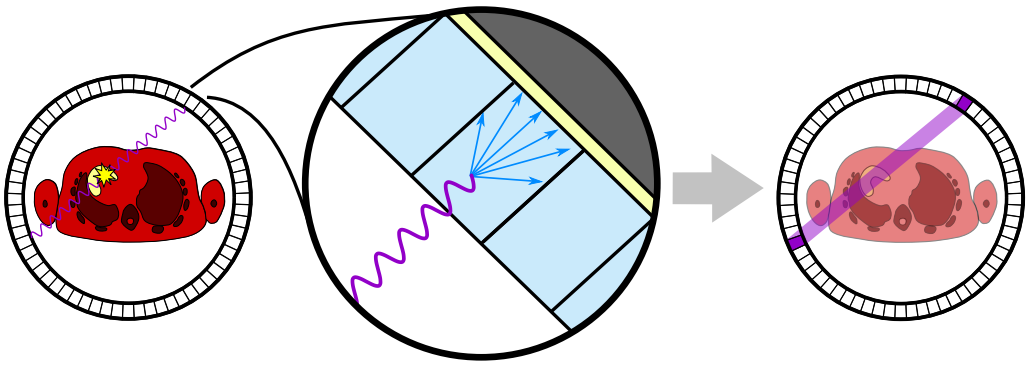

Figure 2.2: **An illustrative example of the PET detection process** This illustration shows a patient surrounded by a ring of detectors (left). The back-to-back annihilation photons can be seen leaving an event in the heart. Each photon can be absorbed by a scintillator crystal, attached to the patient-facing end of each detector. These crystals are designed to scatter an incident annihilation photon, which results in a shower of lower-energy optical photons. These optical photons propagate through the crystal, some scattering to neighbouring crystals or leaving without detection. Many reach the detector end of the crystal to be detected by some photoelectronic device such as a PMT or SiPM (centre), designed to generate a pulse in the electronic circuits. A coincident detection (i.e. detection of two pulses within a short time interval) allows us to infer that the original event took place somewhere within the shaded region (right). Note that the shaded region, or line of response (LOR), encompasses the entire volume between the crystals, not just that between the external faces.

be good at absorbing high-energy photons to increase efficiency of conversion and detection [6].

The light pulse created by the scintillator is then recorded and amplified by the detector, as illustrated in figure 2.2. In a PMT, this is achieved using a circuited window coated with a photoemissive substance, known as a photocathode. The light from the scintillator liberates photoelectrons from the cathode via the photoelectric effect. These photoelectrons are then accelerated down an evacuated tube containing dynodes, which amplify the number of electrons through secondary emission. The amplified signal is then collected at the anode, and converted into an analogue pulse in the electronics.

The amplification factor generated by the dynodes is highly dependent on the voltage applied to the circuit, and is therefore very sensitive to variation [6]. It is for this reason, along with the fact that a current of electrons is heavily affected by a magnetic field, that PMTs are unsuitable for PET-MR applications.

In PET-MR, APDs and SiPMs are used instead. APDs are semiconductor devices which use the photoelectric effect to liberate electrons within the depletion zone of a P-N junction, generating a current. However, they are designed such that many electrons are liberated through secondary ionisation, hence the term 'avalanche'. SiPMs are more modern, and are comprised of many small APDs on a single silicon substrate, which operate in Geiger mode and in parallel [8]. This makes SiPMs generally faster than APDs or PMTs. Both SiPMs and APDs are solid state and therefore stable in a strong magnetic field, making them suitable for PET-MR [8]. The first two commercial simultaneous PET-MR scanners produced to date are the Siemens mMR and the GE Signa PET/MR, which use APDs and SiPMs respectively. GE's implementation of SiPMs also enables its PET scanner to have time-of-flight measurement capability; time of flight allows the scanner to differentiate between the photon arrival times, reducing uncertainty in their origin[1].

As mentioned above, the aim is to identify annihilation events within the FOV. The pulse of current generated by the detector has an amplitude proportional to the energy of the incident photon, allowing the scanner to discriminate between possibly true events and scattered or random events. This is done using a multi-channel analyser (MCA), which discretises the analogue signal. This allows use of a discrete coincidence discriminator, which records an event if two detection pulses occur within an acceptance time. If two signals are recorded within a specified time window (typically about 6–12 ns), they are assumed to have originated from the same annihilation event [6]. Successful detections are recorded with a detection time and the labels of the detectors in coincidence. Further information might also be recorded, such as the energy of the incident photons. A file containing to a list of such recordings is known as a listmode file. Events which register as successful coincidence detections are referred to as 'prompt' events[2]. The rate of detection of single (i.e. not coincident) events can also be recorded for data correction purposes (see section 2.1.4).

[1]The ability to distinguish between photon arrival times allows for further localisation of the point of origin. As a result, the LOR in figure 2.2 would shorten to a 'region of response' instead: this would only incorporate a subsection of the LOR, and with varying probability.

[2]Note that these events are not necessarily 'true' coincidences due to sources of error. This is discussed in section 2.1.4.

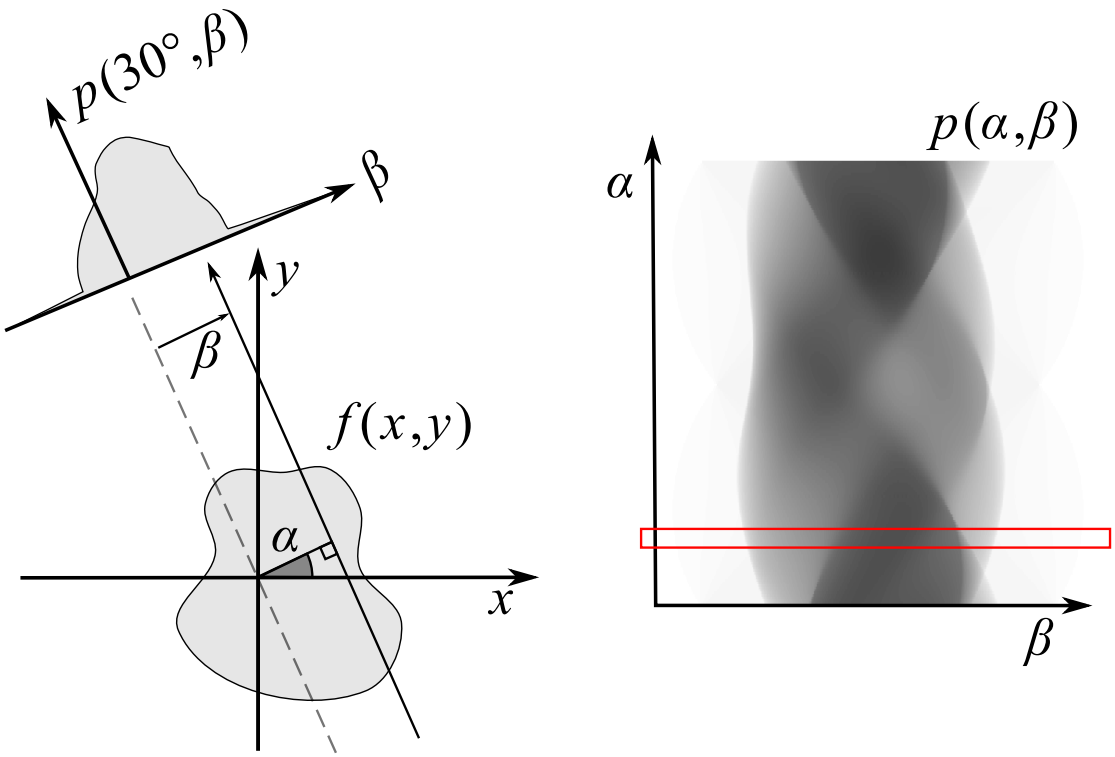

Figure 2.3: **An illustration of the 2D X-ray transform coordinate system** The left hand side of this figure illustrates the geometry of the X-ray transform. The X-ray transform changes the Cartesian coordinate system (x, y) by rotating it by an angle, α. The distribution is then projected parallel to the rotated y-axis, where β describes the signed position away from the origin, parallel to the rotated x-axis. A 1D projection (sometimes referred to as a tomograph), p, can be defined in this way for any angle. The collection of these projections into a 2D coordinate space (α, β) is known as a sinogram, since a point distribution displaced from the origin will trace a sinusoidal path through the sinogram. The distribution shown on the left produces the sinogram on the right, where the location of the example projection is indicated by a box over the sinogram. PET sinograms are conventionally displayed on the interval $[\,0°, 180°)$.

2.1.3 Sinograms and Scanner Geometry

As described above, coincident events can be recorded in a chronological list, known as acquiring in list mode. This can be particularly useful for dynamic studies, or applications requiring a time component such as motion correction [6]. However, recording information about every detected coincidence can result in the listmode file being impractically large. The number of possible combinations of all crystals in a 3D PET scan is also very large, leading to high computational costs during reconstruction.

This can be addressed by compressing the data into the image-like sinogram format. The word 'sinogram' refers to the sinusoidal appearance of objects in the sinogram, as shown in figure 2.3. The listmode data are converted into this format by binning the events into their respective detector pairs, akin to a histogram. In 2D PET scans, sinograms have two axes, corresponding to the combinations of these pairs. The pairs are ordered such that each row can be considered a 1D projection of the FOV at a given angle, with the angle varying between 0° and 180° along subsequent rows. In 3D scans, sinograms can have two additional axes, corresponding to inter-ring coincidences being organised into all pairs at a given oblique angle, for all possible angles. Following the histogram analogy, this is where compression is possible: these axial pairs can be grouped to reduce the number of combinations that need to be calculated when processing the data. This is described in more depth in figure 2.4.

The mathematical description of the correspondence between 2D real space, (x, y), and sinogram space, (α, β), is known as the X-ray transform,

$$\mathcal{R}\{f(x,y)\} \equiv p(\alpha,\beta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x\cos\alpha + y\sin\alpha - \beta) f(x,y)\mathrm{d}x\,\mathrm{d}y \qquad (2.1)$$

where $p(\alpha, \beta)$ is the projection of activity distribution $f(x, y)$ observed at angle α and normal offset β. This coordinate transformation is illustrated in figure 2.3. The coordinate system (α, β) is precisely that used to order the detector pairs in a sinogram; each row is a 1D projection of the FOV at a given angle, α.

In 2D PET, one sinogram is generated for each axial ring of detectors. The advancement of 3D PET dramatically increased the number of possible LORs across the FOV, largely increasing the number of 2D sinograms that can be generated. The third dimension of sinogram data is somewhat convoluted: as before, it contains the in-plane sinograms (one for each ring), but once all these are exhausted, there are further sinograms corresponding to oblique rays across the axial rings. The specifics of the ordering of these oblique sinograms comes with its own terminology, and is often simplified by representing the cross-ring combinations and their compression using a 'Michelogram', as shown in figure 2.4 [9].

Finally, it is worth noting that listmode data are recorded to a precision of one mil-

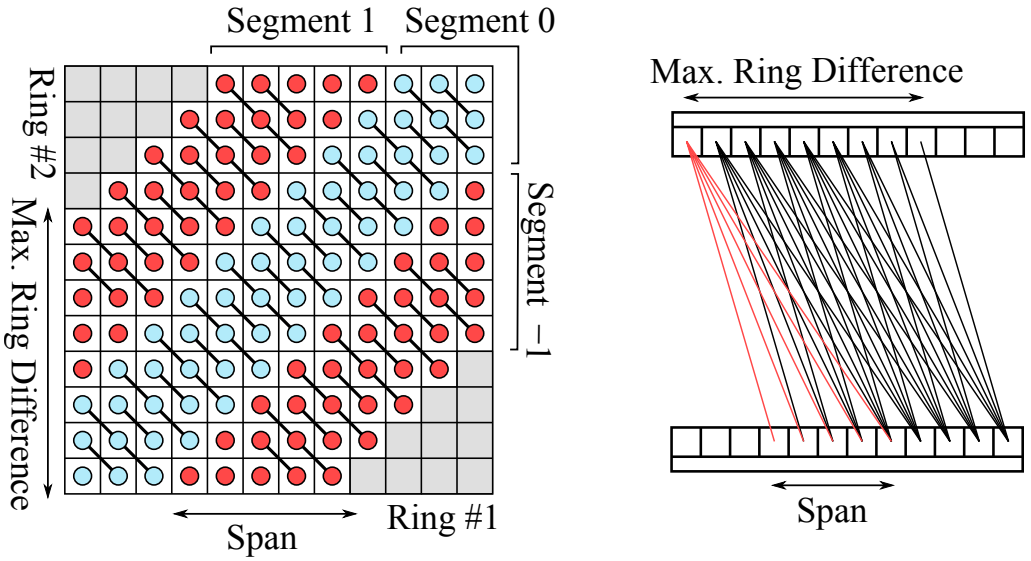

Figure 2.4: **Nomenclature of data compression in 3D PET** In 3D PET, the number of possible crystal combinations between axial rings drastically increases. The graphical representation of axial positions of 2D sinograms is known as a Michelogram (left). These can be used to show how 2D planes can be combined to optimise storage space using axial compression [9]. This example is a 12 ring scanner. Each bin on the graph represents a 2D sinogram, with the axes representing the crystal ring combinations. Sinograms up the leading diagonal represent sinograms in the axial plane. Off-diagonal elements represent cross-ring sinograms. Sinograms with close axial positions can be grouped into 'segments' [9]. Sinograms within a segment can be combined to form a new Michelogram of fewer combinations. However, this causes losses in axial resolution. The number of sinograms combined across rings either side of an axial plane is called the span [9], which in this case is 5. The leading diagonal segment is labelled segment 0, with successive segments either side being labelled as ± 1, ± 2, etc. Sinograms with a high ring difference will have longer LORs through the rings prior to detection, degrading resolution. A maximum ring difference can be applied, which in this example is 8. Sinograms beyond this can be discarded to further reduce storage requirements, as indicated by the greyed-out bins. An illustrative cross-section through the rings of the scanner is shown (right). The lines represent all possible planes in segment 1. As indicated by the planes highlighted in red, each ring connects to a maximum of 5 others, as dictated by the span. Notice that planes in successive segments become increasingly oblique, and therefore have a longer path length.

lisecond on most commercial scanners (including the Siemens mMR), which allows for excellent time resolution in the data. This allows data to be binned along an extra dimension in sinogram space, which typically occurs in two ways. The first is referred to as dynamic binning, where each set of sinograms corresponds to a unique sub-interval of the total scan time. This is particularly useful for kinetic studies, since each subset can be reconstructed into an image with a time-dependent distribution of radiotracer. The second method is known as gating, in which the data are binned into multiple sinograms according to a specified rule, which is generally periodic in time, such as cardiac phase (an illustration of this can be seen later, in figure 4.1). Gating is useful for motion correction, since gates covering short sub-intervals in a period will contain a smaller range of motion positions, resulting in less blur. In general, a gate will accumulate many more counts over a scan than a single dynamic frame. These topics are discussed in more detail in chapter 4, and gating is discussed in section 4.2.1.

2.1.4 Sources of Systematic Error

The use of high-energy photons to record positron annihilation events can be challenging. Some of these challenges are explained in figure 2.5. Firstly, ethical and practical considerations limit the amount of activity that can be used. The use of relatively low activity results in annihilations and their subsequent detection being relatively rare. Single events, due to one of the coincident pair being undetected or due to background radiation sources, are much more common, such that the scanner will typically only convert 1% to 10% of single events into a detected coincidence (which itself may or may not be a true coincidence, explained below) [9]. Accidental coincidences may also occur due to the presence of background radiation.

Secondly, matter-photon interactions can cause systematic errors when trying to identify true coincidences. For example, materials with a high atomic number will have a greater stopping power for deflecting or absorbing the annihilation photons. In turn, this leads to a loss of signal along paths through dense objects, or deflection of annihilation photons off their original course, leading to misidentification of the line of origin.

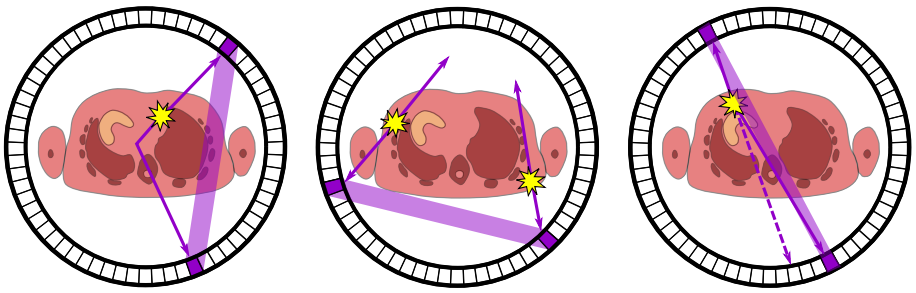

Figure 2.5: **Sources of error in PET data** This figure shows 3 examples of the causes of error described in section 2.1.4. Scattering of one (or both) incident photons can lead to an angular departure from the true LOR (left). Random events, such as misidentifying two unrelated photons as originating from the same event, can affect signal (centre). A small but non-zero momentum remaining when the positron and electron annihilate can lead to a slight angular discrepancy in the linearity of the photon paths. Although exaggerated here, this effect becomes more significant with larger detectors, and contributes to the effective scanner resolution, due to slight LOR mismatching (right). The misattributed LORs are illustrated for each situation: notice that these do not, in general, contain the event(s) which produced the detected photons.

This brief list is not exhaustive; there are many additional considerations such as geometric effects, electronics design, normalisation of detectors, and so on. To obtain high-quality quantification in PET images, it is important to identify these errors and attempt to correct for them. However, this is not always straightforward. The following section will describe the most significant sources of error (except for motion; as the focus of this thesis, this will be discussed separately in the next chapter), and how we generate estimates of these contributions to the measured data.

Attenuation

Attenuation can be a major cause of signal loss. Attenuation occurs due to photon-electron interactions within a medium, with higher electron densities and atomic number causing more absorption and scattering. The fraction of photons not absorbed or scattered can be calculated according to

$$T = \exp\left(-\int_0^{\ell} \mu\left(z\right) \mathrm{d}z\right), \tag{2.2}$$

where T is the proportion of photons transmitted after travelling along total path length ℓ through attenuating material(s) with linear attenuation coefficient $\mu(z)$,

usually quoted in cm^{-1}. In PET applications, this path is a straight line. Since both photons are back-to-back, it can be shown that the total attenuation along an LOR is independent of the origin of the photons along that line, such that the probability of both photons being detected depends upon the length of the line and the sum of all attenuation coefficients along it.

The total attenuation caused by an object can be measured directly with a transmission scan, which involves acquiring from a rotating source of 511 keV photons, such as germanium-68. This is compared to a blank scan to calculate the percentage drop of transmitted photons when the subject is in the FOV. As described above, the probability of detecting a photon pair after passing through an attenuating material is independent of their origin. This can be used to approximate the percentage drop in the i^{th} sinogram element as

$$a_i = \exp\left(-\alpha_i\right) \tag{2.3}$$

where α_i is the total linear attenuation coefficient along the straight line between the i^{th} pair of detectors. The set of values $\boldsymbol{a}$ in (2.3) is occasionally referred to as an attenuation-factor (AF) sinogram, since it contains values ranging continuously between 0 and 1, corresponding to the average drop in number of coincidences due to attenuated photons. The transmission scan method described above directly generates an AF sinogram, although an image of μ values, known as a μ-map, can be produced by reconstructing the AF sinogram.

Modern clinical PET scanners tend to use CT instead, since it is faster and dual modality PET-CT scanners are currently the clinical norm. This uses a polychromatic X-ray source to calculate the electron density of tissues in the subject as a 3D image. These are generated in Hounsfield units, which are standardised against water. The image produced by CT is a type of μ-map, but measured around the 30 keV to 100 keV energy scale. Direct conversion between this energy range and the 511 keV scale required for PET is not straightforward because μ in (2.2) is both energy- and atomic number dependent. This is sometimes addressed by segmenting the tissues in a CT image (into regions such as soft tissue, cortical bone, and lung tissue) and assigning them known μ values from the gold-standard set by transmission

scans, forming a CT-derived μ-map [6]. Otherwise, a bilinear relationship is used to map between Housfield units and the linear attenuation coefficient at 511 keV, with an inflection at the origin.

Since MRI measures the spatial distribution of hydrogen nuclei in the subject, there is no clear way to scale between an MR image and the electron densities required for a PET μ-map. Current methods involve segmentation of MR images, often from Dixon scans, [10, 11], which can work well for soft tissues. However, the signal produced by cortical bone (a major contributor to local attenuation effects) is too fast for the receiver coils to detect, and as such bone gives little to no signal in the final image. There is active research trying to improve the signal gained from bone, either directly or through inference, using ultra-short echo time (UTE) scans [12]. At the time of writing, reliable measurement of attenuation in whole-body simultaneous PET-MR remains an unsolved problem. Recent work by [13] summarises the recent advances in this field.

Scattering

Compton scattering can lead to one or both coincident annihilation photons to be scattered prior to detection [9, 14]. This leads to loss in photon energy and an angular deflection, which in turn causes the coincidence to be detected on a different LOR than the original one [9]. The amount of energy lost is dependent upon the angle of deflection, with 50% of all Compton interactions scattering at an angle lower than 60° or less [14].

An example of a scattered coincidence is illustrated in figure 2.5. The misidentification of the true LOR causes a broadening of the 1D projection distribution at each detection angle in the prompt sinogram, reducing the contrast in the final image [6].

Energy discriminators in the MCA circuit are designed to lower this, but photons scattered by a small angle retain much of their energy, the detection of which is dependent on the energy resolution of the detectors. To some degree, the lead septa in 2D PET scanners mitigated some of the scatter effects. The problem became much more serious with the introduction of 3D PET.

Since detection can be difficult, scattered events are instead modelled to estimate their contribution to the data. Monte Carlo simulations can use the underlying physics to predict a scatter distribution for the purposes of correction, although this can be resource-intensive [14]. Alternatively, one can use approximate analytical models [14], or the single scatter model, which can efficiently estimate the scatter distribution under the simplifying assumption that the photons can only scatter once, known as the single-scatter simulation (SSS) [15].

Random Events

The finite width of the timing window when searching for events can lead to detection of random, or accidental, coincidences. These can include apparent detection of coincident photons that occurred from different mother nuclei, accidental coincidence with a background gamma ray, and even detection of the scintillator's own radiation (this last point occurs because inorganic scintillators are often comprised of heavy elements such as lutetium) [9]. An example of uncorrelated photons being detected is shown in the centre of figure 2.5.

Random events tend to be distributed uniformly across the FOV, leading to an overall reduction of image contrast if left uncorrected. The randoms, R, detected along LOR (i, j) can be modelled by the equation

$$R_{ij} = 2\tau S_i S_j \tag{2.4}$$

where S_i and S_j are the rates of single events detected by detectors i and j respectively, and τ is the width of the timing window [9].

Randoms can also be estimated using delayed coincidence detection. This relies on the use of two coincidence circuits, but with offset timing windows. The first is the regular prompt circuit, and the second is the delayed circuit. The delayed circuit can have a longer τ, which detects delayed coincidences. These delayed detections are assumed to only consist of randoms [16, 6].

Randoms, measured or modelled, can be subtracted from prompt data as a cor-

rective measure. However, this can result in an increase in statistical noise in the reconstructed image since the variance of the noise quality of the data is no longer Poisson, as is assumed in many iterative reconstruction techniques [16]. Instead, modern reconstruction techniques include these effects in a model of the data when reconstructing the image, compensating for the errors without the noise amplification.

Normalisation

A PET scanner is composed of many detectors working in unison. Normalisation is the process of standardising the output of each detector [14]. This can account for geometric effects, such as the variation in LOR width due to crystal depth (examples of this can be seen in figures 2.2 and 2.5: the endpoints of each LOR are not constrained to the inward-facing crystal faces). It can also account for practical sources of error, such as the variation in relative sensitivity of each detector. Normalisation is performed using a flood sinogram, which is created by exposing all detectors to a uniform activity distribution (such as a rotating rod source of germanium-68). To avoid issues with deadtime (below), this activity source must be relatively weak, meaning that these scans can take a long time to acquire good SNR.

Deadtime

An important practical issue during a scan is deadtime: time lost during the scan, leading to low-efficiency scanning. Several things can cause this. Firstly, each scintillator crystal can only report one event between the absorption of the annihilation photon and the subsequent decay that produces optical photons. Any other annihilation detections in the same crystal during this time will not be measured. Secondly, the digital electronics only carry pulses up to a certain rate. If two detections are made within the same energy window, they may be characterised by the energy discriminator as a single, large-energy event, and thus ignored as a false signal. Usually events are sufficiently infrequent that deadtime is not a problem.

However, it can be an important consideration for scans involving very high rates of activity [6].

Resolution Effects

A number of physical effects contribute to the intrinsic resolution of PET scanners, often referred to as the 'point-spread function' (PSF) of the scanner. Firstly, the angular size of the individual detector elements, d, leads to a varying resolution in the direction parallel to the detector face, ranging between $d/2$ at the centre of the FOV and d at the detector itself [6]. Secondly, it is usually assumed that the annihilation photons are emitted exactly 180° apart. In practice this is rarely true, since there is often residual energy in the positronium system at annihilation, leading to an angular discrepancy of approximately 0.5°. The impact of this deviation depends on the distance between the two detectors used to detect it, D, such that $R_{\mathrm{non-colin}} = 0.0022D$. For a typical whole-body PET scanner of diameter 80 cm, this leads to a maximum contribution of 2 mm [6]. An example of the non-colinearity of photons is illustrated in the right-hand-side of figure 2.5. Another contribution to intrinsic resolution is the effective distance the positron propagates prior to annihilation. This ultimately depends upon the initial energy of the positron (the maximum of which varies with the mother isotope), but for fluorine-18 in water, the effective range is approximately 0.3 mm [6].

The contributions of each of these effects to the scanner resolution can be calculated by adding the respective resolutions in quadrature [6]. Overall, these effects combine to a fundamental physical limit for the system resolution[3] between a few tenths of a millimetre to a few millimetres [6]. Point-spread function modelling, and how to most effectively incorporate the point-spread function into the image reconstruction, is an active area of research in an attempt to improve resolution. For further information, please refer to [17].

Related to the above is the partial volume effect, which depends on the relative contrast and distribution of activity. Due to the image sampling, there can be a re-

[3]System resolution is a fundamental limit due to the physics of the processes involved. It should not be confused with image resolution, which may vary due to data processing.

duction in the apparent activity of small features (smaller than approximately twice the system resolution), which may also spill over into neighbouring regions, increasing the background apparent activity. The fact that this happens in both directions (background to feature, and feature to background) makes disentanglement hard to address [18].

Finally, the ability to resolve small structures can also be limited by motion. This is discussed at length in chapter 3. Benefits from improvements in resolution can be limited until impacts of motion can be minimised [19, 20].

2.2 Image Reconstruction

The ultimate aim of medical imaging is to obtain useful information about the subject to aid in disease detection, staging, or assessing efficacy of disease treatment. In PET, which is often used as a quantitative tool both clinically and academically, it is important to establish reliable methods of image reconstruction. Noise levels in PET are very high relative to the signal. Reconstruction techniques therefore need to be reliable, accurate, and robust to high levels of noise.

2.2.1 Analytical Reconstruction

As mentioned in section 2.1.3, data can be stored as a set of sinograms. Sinograms are useful because they set the data into a geometry described by the X-ray transform, described by (2.1). This allows us to use the mathematics of the integral transform to invert the transformation back into Cartesian space. Analytical reconstruction methods tend to be derived using a continuous formulation of images and data.

Inversion of the X-ray transform involves the concept of back-projection. The back-projection (BP) transformation is defined by

$$\mathcal{R}^{\dagger}\left\{p\left(\alpha,\beta\right)\right\} \equiv \hat{f}(x,y) = \int_{0}^{\pi} p\left(\alpha, x\cos\alpha + y\sin\alpha\right) \mathrm{d}\alpha. \tag{2.5}$$

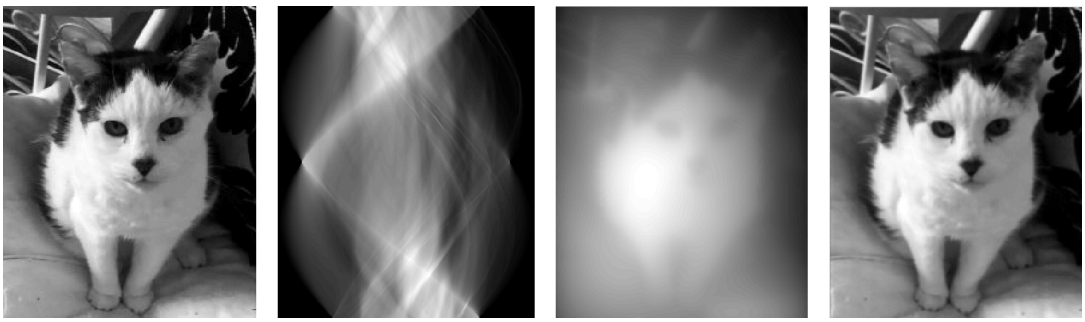

Figure 2.6: **Back-projection vs. filtered back-projection** The principles discussed in this section are not specific to medical imaging: The X-ray transform can be applied to any image. In this figure, a 2D image of a cat has been transformed into a sinogram. As explained in figure 2.3, each row of a sinogram is the 1D projection of the 2D distribution at a given angle. To recreate the image through analytical reconstruction, each of these projections can be projected back through image space. This process, known as back-projection, recovers a blurred version of the original. To obtain a more faithful image, the blur must be accounted for using filtering methods, a process known as filtered back-projection (FBP), shown on the far right. However, this does not cope well with noise in the sinogram. An example of this is shown in figure 2.7.

This integral transform takes a projection, p, and projects it back across the FOV along all parallel lines at the angle α [6]. The need to do this is intuitive: when applying X-ray transform originally, the exact contribution of each point on a line to the resulting projection distribution is lost due to superposition. Therefore, without prior information on the locations of these contributions, the next best guess is to assume each point in the image contributed equally to the projection measured along a given LOR.

Note that simply applying BP will not invert the X-ray transform alone: back-projecting each projection across Cartesian space at its respective angle does not yield the original distribution, f; instead we yield the back-projection, f_{BP}. This distribution appears to be a blurred version of f because the coordinate change from radial to linear coordinates leaves behind a radial dependence in k-space. An example of this can be seen in figure 2.6. The relationship between f_{BP} and f is described by

$$f_{\mathrm{BP}}(x, y) = \frac{1}{\sqrt{x^2 + y^2}} \star f(x, y) \tag{2.6}$$

where $\star$ is the convolution operator [6].

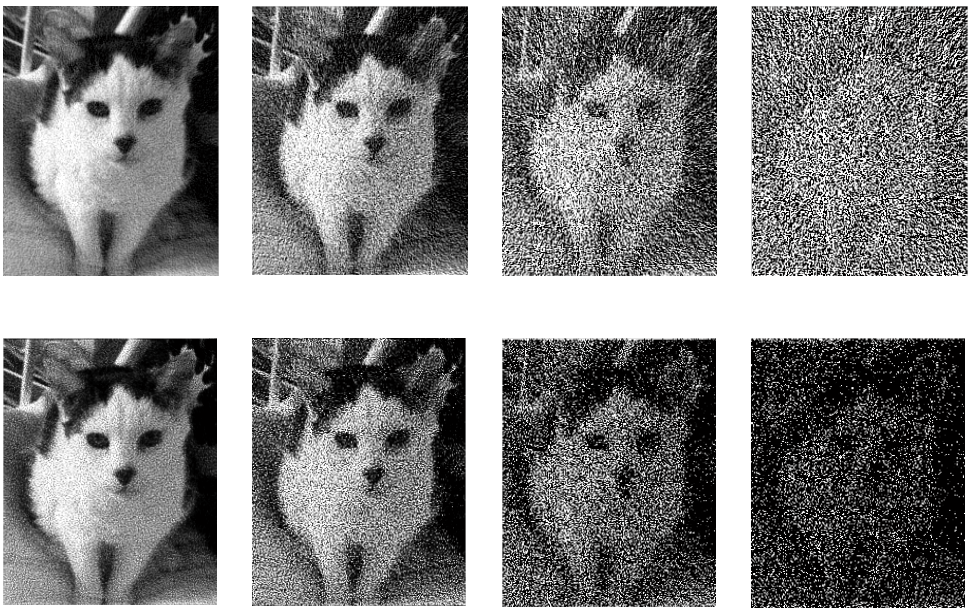

Figure 2.7: **Examples of FBP and MLEM with increasing levels of noise** FBP reconstruction does not cope well as the noise level in the sinogram increases. The top row of images is comprised of increasingly noisy sinograms reconstructed with FBP, using a ramp filter. Reconstruction of the same sinograms using MLEM (1000 iterations) results in the images in the bottom row, which in the noisiest case can recover a more recognisable image. Note that all images are shown with the same intensity scale.

Direct Fourier Reconstruction

A property of the Fourier transform can be used to circumvent the time-consuming BP step during inversion. This property is known as projection-slice theorem. The projection-slice theorem shows that the Fourier transform of a projection at angle α is equivalent to a 1D sample of the 2D Fourier transform of the whole object in k-space at the same angle [6]. This method allowed reconstruction without the use of filtering. However, since the Fourier data was constructed in polar coordinates, this introduced issues with sampling back to a Cartesian grid, which was computationally intensive and prone to creating image artefacts.

Filtered Back-Projection Reconstruction

Another analytical reconstruction algorithm is filtered back-projection (FBP). This combines the use of the BP operator and projection-slice theorem to include a filtering step, whilst avoiding the change between polar and Cartesian coordinates

that occurs in direct Fourier reconstruction [6]. This time, each 1D projection is convolved with a filter, such as $|k|$, prior to being projected across Cartesian space. This addresses the blurring observed when using regular back projection, as can be seen in figure 2.6.

The introduction of 3D and dynamic PET posed a serious problem for FBP, since reconstruction times took much longer. Additionally, some of the acquired projections in 3D PET did not contain sufficient information to reconstruct without introducing artefacts. Attempts to improve included the 3D reprojection (3DRP) method, which used 2 passes of FBP to estimate missing data [21]. However, this is computationally demanding, and could take up to 30 times longer than 2D FBP [22]. This was improved by the development of 3D to 2D rebinning algorithms, such as Fourier rebinning (FORE) [23], which converted 3D datasets into a set of multiple 2D sinograms.

Note that the convolution step in FBP supresses low-frequency components in k-space and amplifies high frequencies. This has the effect of sharpening edges (which can lead to ringing artefacts) and increased levels of noise in the image. In fact, FBP doesn't cope well with low SNR in the data at all. Examples of this can be seen in figure 2.7. This problem is exacerbated by attempts to remove systematic errors, through (for example) attenuation correction and randoms correction, which can increase the variance in the data [16]. This can be somewhat improved by using filters modulated at high frequency, such as the Hann filter. Over time, iterative reconstruction techniques were developed which recovered more image contrast in low-count situations and allowed modelling of error effects. This is now generally the preferred class of reconstruction methods.

2.2.2 Mathematical Descriptions of Digital Images

Before continuing the discussion of PET reconstruction, it is important to establish what we really mean when we discuss images in a mathematical sense. The lay-definition of an image is often synonymous with a picture or photograph. This is also somewhat true in research fields involving imaging, although in this case we must

be more careful with the technical aspects of how an image is said to 'represent' the object. For example, consider a photograph: it is generally regarded as a snapshot of the subject. Technically, the photograph is a 2D view of a 3D distribution of objects, which is limited by things such as the quality of the lens, the size of the aperture, and the resolution of the detectors (i.e. film or charge-coupled device) used to record the image. This has much in common with the previous section on the practicalities of PET imaging. It is especially important when deciding how to convert data acquired from an imaging device into the final image.

The lay-definition also tends to consider the object in an image as a continuous object. Again, in a technical sense this is not strictly true: storage of this information is discretised. A photograph appears continuous because the human eye cannot resolve the silver halide crystals used to store the image. Similarly, digital images are often comprised of a finite number of discrete elements such as pixels. If an imaging system, such as the PET scanner and its geometry, are to be modelled in a continuous way (e.g. for modelling motion), the relationship between the continuous and the discrete must be considered. The formalities of this will be described below.

Discretisation

As mentioned in the introduction section, the technical definition of an image is required to make sense of the data reconstruction problem. Due to the nature of the data being acquired, this thesis will focus on greyscale digital images. Mathematically, one may think of a grayscale image as a representation of a scalar field, f. A scalar field is a function which has a single value at every point in N-dimensional space[4], denoted $\vec{r} \in \mathbb{R}^N$.

Often, images are simply considered as an ordered set of J intensities, $\{f_1, f_2, \ldots, f_J\}$, arranged on an N-dimensional regular grid of known size, $[n_1, n_2, \ldots, n_N]$. Sometimes, these f_j are referred to as samples of f. In the description of an algorithm, the image can be expressed as a single vector (specifically, a $[J \times 1]$ data vector),

[4]This ambiguity in the number of dimensions is deliberately maintained since the definitions in this section hold for images in 2, 3, and even 4 dimensions. This 4-dimensional case is 3D plus time.

denoted $\boldsymbol{f}$. Note the use of different notation for spatial vectors, such as $\vec{r}$, versus data vectors, such as $\boldsymbol{f}$. This is both for readability, and to avoid any ambiguity in later discussions in this thesis. This distinction is particularly useful when there is a corresponding data vectors for each of Cartesian dimension (i.e., a list of 3-dimensional grid points, $\vec{\boldsymbol{r}}$).

The set of intensities as a data vector is usually sufficient for intuitive understanding of many algorithms in imaging science. However, it is possible to take this further: the grid on which the intensities sit can be included as part of the image definition. This enables direct mapping from the continuous function that the image seeks to represent, f, to the intensities, $\boldsymbol{f}$. This is done by introducing 'basis functions', $b = b(\vec{r})$, which relate the discrete elements of $\boldsymbol{f}$ to the space on which they are supported, $\mathbb{R}^N$ [24, 25]. Generally, a basis function can be defined for each intensity, forming a set of them $\{b_1, b_2, \ldots, b_J\}$. Now we can approximately equate the image to the original field according to

$$f(\vec{r}) \sim \sum_{j=1}^{J} f_j b_j(\vec{r}). \tag{2.7}$$

Note that (2.7) can become an exact equality in only certain cases, and other formulations are possible.

A common example would be to consider a 2D greyscale image on a computer screen, which is comprised of pixels. Pixels are squares, arranged at regular intervals, each with its own intensity value. The pragmatism behind choosing pixels is clear: the square tessellates, leading to good packing without any gaps. From a technical standpoint, however, each pixel has a *spatial* extent too: an object that is 10 pixels across in the image might be 1 cm in reality. Thus, each pixel has a single value of intensity 'supported' across the millimetre it represents; any information about the variation of f within the boundaries of this pixel has been lost. In reality, f may vary smoothly between these points, and it is due to this that (2.7) is only an approximation.

Examples of Image Basis Functions

There are many possible choices for basis functions in (2.7), depending upon the application. In the example we discussed the pixel, also known as a voxel (3D), or a doxel (4D). Defining b_{pixel} relies upon the 1D Heaviside step function, where

$$H(x) = \begin{cases} 0 & \text{if} \quad x < 0 \\ 1 & \text{if} \quad x \geq 0 \end{cases} \tag{2.8}$$

which can be used to construct the top-hat function, $\Pi(x) = H\left(x + \frac{1}{2}\right) - H\left(x - \frac{1}{2}\right)$. An N-dimensional analogue of this top-hat can be created through a product of N multiples of Π, each taking its respective ordinate (x, y, and so on) as the argument. Thus, 2D images with pixels a set of shifted basis functions, where $b(x, y) = \Pi(x)\,\Pi(y)$ is the basis function of a single pixel at $(0, 0)$. The exact dimensionality of the function isn't of consequence in this thesis; $b_{\text{pixel}}(\vec{r}) = \Pi(\vec{r})$ will be used in general.

As mentioned, pixels are useful due to their ability to tessellate. This can minimise any processing required to display images. However, the pixel might not always be the ideal basis function for the problem at hand. For example, it might be favourable to choose to represent the image using overlapping, truncated Gaussians (known as 'blobs', sometimes also modelled using Bessel functions) which can be more practical for modelling the PSF in PET reconstruction [24, 26].

Another interesting example is the use of box splines, which have closure under the X-ray transform. This means that the X-ray transform of a box spline is always itself a box spline [25]. The same cannot be said for pixels since, for example, the projection of a square at any angle in the interval $[0°, 90°)$ is itself *not* a square. Thus pixels are not closed under the X-ray transform and care be taken to consider the effects of this change of basis when discussing X-ray transforms applied to a pixel-based image.

Basis functions are discussed here because they will become relevant when defining the motion estimation problem in chapter 7. Although this thesis only employs a pixel basis, attempts have been made to maintain some level of generality in

discussions later on.

2.2.3 Iterative Reconstruction

Enabled by improving processing power, analytical reconstruction methods have slowly been phased out in favour of iterative approaches. These iterative methods cope better with data imperfections such as low noise or, for example, attenuation effects through statistical modelling. This has become increasingly important for more contemporary research, such as dynamic PET or motion correction, because they generally require separation of the data into frames (for dynamic studies) or gates (for motion correction), which further reduces SNR. Iterative reconstruction is thought to generally provide a better solution (for example, in image quality) than analytical methods [16], although iterative methods might not be the best choice for low-count datasets requiring good quantification, such as in dynamic PET studies. Iterative reconstruction is the method used in this thesis, and the following contains more detail on the iterative reconstruction methodology.

Generally iterative approaches involve an initial estimation of the activity distribution, which is then repeatedly updated and improved until a solution is found. Iterative reconstruction algorithms have three major parts:

1. **Modelling the data:** A mathematical relationship is formulated, formalising the link between the true spatial distribution of the radiotracer and its appearance in the PET data.

2. **Creating an objective to utilise this model:** An objective function is devised, which can be used to compare the measured data to the data predicted by the model. This allows quantification of success for a given image estimate.

3. **Finding a means to estimate the best image, according to the objective:** The objective function is used to successively update an estimate of the tracer distribution according to some optimisation algorithm.

There is a large body of iterative reconstruction algorithms in the literature, but

generally all tend to follow this pattern. The following subsections will explain these components of the reconstruction in more depth.

Modelling the PET system

Accurate modelling is important since it is used to compare an estimate of the activity distribution to the real, acquired data. Failure to model the measured distribution accurately can lead to some sources of error being uncorrected in the final image, or might even introduce additional systematic errors such as artefacts. In many cases, the data modelling is probability based: the acquired data are considered a single noise realisation of an underlying mean spatial distribution, with characterisable statistics. This allows noise handling through incorporation of a statistical distribution, which in PET is often the Poisson distribution.

In sections 2.1.3 and 2.2.1, the X-ray transform and back-projection were presented as continuous transformations. For convenience, the X-ray transform in (2.1) is repeated here:

$$\mathcal{R}\left\{f(x,y)\right\} \equiv p(\alpha,\beta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \delta(x\cos\alpha + y\sin\alpha - \beta) f(x,y)\mathrm{d}x\,\mathrm{d}y.$$

Digital images, on the other hand, are stored as discrete objects. The correspondence between the projection space, $\vec{u}$, and Cartesian space, $\vec{r}$, was introduced in section 2.2.2. It is possible to express the X-ray transform in matrix form:

$$[\boldsymbol{A}]_{ij} = \iint K(\vec{r},\vec{u})d_i(\vec{u})b_j(\vec{r})\mathrm{d}\vec{r}\,\mathrm{d}\vec{u}, \tag{2.9}$$

where K is a generalisation of the kernel in (2.1), and $\{d_i(\vec{u})\}$ and $\{b_j(\vec{r})\}$ are the sets of basis functions for sinograms and images, respectively. This matrix, $\boldsymbol{A}$, is often referred to as the 'system matrix' in the PET literature. Provided the sinogram basis is orthonormal[5], (2.1) can be written as a system of linear equations,

$$\boldsymbol{q} = \boldsymbol{A}\boldsymbol{f} \tag{2.10}$$

[5]Failure of the orthonormality condition results in the equation $\boldsymbol{O}\boldsymbol{q} = \boldsymbol{A}\boldsymbol{f}$ instead of (2.10), where matrix $\boldsymbol{O}$ accounts for any overlaps in the sinogram basis functions, $\{d_i\}$.

where $\boldsymbol{f}$ are the coefficients of the image basis functions for the image estimate, and $\boldsymbol{q}$ the coefficients of the sinogram basis functions. These are $[J \times 1]$ and $[I \times 1]$ vectors, respectively, where I is the number of sinogram elements.

The exact form of the system matrix depends on the geometric relationship between an image of the radiotracer distribution and its appearance to the scanner. In probability-based algorithms, the system matrix models the mean measurement of the scanner. In these cases, each element of $[\boldsymbol{A}]_{ij}$ represents the probability of detector pair i detecting a coincidence from an event in pixel j. Methods of incorporating additional effects in the data into the system model will be discussed in later chapters.

Equation (2.10) appears relatively simple. However, both I and J can be upwards of 10^4, such that there are more than a billion elements in $\boldsymbol{A}$. This makes storage and subsequent inversion of the system matrix impractical. Instead, iterative methods are required, using a series of forward-projection and back-projection operations ($\boldsymbol{A}$ and $\boldsymbol{A}^{\mathrm{T}}$ respectively) to tend towards a reliable solution for $\boldsymbol{f}$.

Objective Functions

All iterative algorithms work to minimise or maximise some measure of success, known as a cost or objective function. This is done by making an initial estimate of the image, denoted $\boldsymbol{f}^{(0)}$. The data model is then used to generate the sinograms that would be expected if the tracer distribution was as expected in the image estimate. The objective function then compares this data estimate with the measured data, and provide an update to the initial image, to obtain $\boldsymbol{f}^{(1)}$. Methods for the update step are discussed in the next subsection. The updated image is then used as the next image estimate, which is updated, and so on. Iterations will be denoted using a bracketed superscript index k, such that the k^{th} estimate of $\boldsymbol{f}^{\text{true}}$ is $\boldsymbol{f}^{(k)}$.

Objective functions can be straightforward, such as the least-squares objective:

$$Q_{\text{LS}}\left(\boldsymbol{f}^{(k)}\right) = \sum_{i=1}^{I} \left(m_i - q_i^{(k)}\right)^2. \tag{2.11}$$

This compares all projections of the measured data, m_i, with the equivalent projection of the current model of the data, $q_i^{(k)}$, which is a function of the image estimate, $\boldsymbol{f}$. Taking the square of this difference provides stronger penalisation for estimated values that deviate further from the data. This method does not cope well with noisy data [16].

To address the noise issues affecting (2.11), a weighted least squares objective function was designed. This incorporates a statistical weighting, σ_i, to each sinogram pixel:

$$Q_{\mathrm{WLS}}\left(\boldsymbol{f}^{(k)}\right) = \sum_{i=1}^{I} \frac{\left(m_i - q_i^{(k)}\right)^2}{\sigma_i^2}. \tag{2.12}$$

This weight is typically derived from an assumption that the noise is Gaussian in nature [16]. Since the noise in raw PET data is Poisson, this approximation can be a good one if there is a high number of counts in each projection bin, q_i.

Once a suitable objective function is designed, an optimisation algorithm is employed. For example, Q_{LS} should minimise as $\boldsymbol{f}^{(k)}$ approaches the solution, $\boldsymbol{f}^{\mathrm{LS}}$. This minimum can be found by several methods, such as gradient descent, which takes the gradient of Q_{LS} and uses it to converge towards $\boldsymbol{f}^{\mathrm{LS}}$.

Optimisation Schemes

As mentioned above, the objective function quantifies the success of the image estimate. Generally speaking, no estimate is correct on the first guess, and thus a method to improve the estimate is required. This is achieved using optimisation methods.

The simplest and most straightforward method of optimisation would be an exhaustive search. This involves randomly trialling many values for the estimated parameter and testing the success of each one in turn using the objective function. Although easy to implement, stable, and reliable, the high computation costs required for multivariate systems (such as a PET image with many intensity values to estimate) make it highly impractical.

Another set of optimisation methods is gradient-based algorithms. These use the

derivative of the objective function to infer the direction in which the solution can be found. The simplest method is gradient ascent (or descent, depending on the objective used), for which

$$\boldsymbol{f}^{(k+1)} = \boldsymbol{f}^{(k)} + \beta^{(k)}\boldsymbol{g}^{(k)}, \tag{2.13}$$

where k records the iteration number, $\boldsymbol{f}$ the image intensities, $\boldsymbol{g}$ the gradient of the objective for each intensity, and β a stepsize to stabilise the optimisation and encourage convergence. Poor choice of stepsize or objective function can render this method ineffective, since local extrema may exist. Pre-conditioning can be used to improve the convergence properties of gradient ascent, especially when combined with conjugate-gradient algorithms [16].

The issue of convergence was addressed by the development of the maximum-likelihood–expectation-maximisation (MLEM) algorithm, which provided a well-defined stepsize from its formulation, applied as a multiplicative update which is guaranteed to converge to the global maximum of the Poisson objective function. The derivation of the MLEM algorithm will be presented in the following subsections.

The Poisson Log-Likelihood Objective Function

The MLEM algorithm was introduced by Dempster in 1977 [27] and developed for PET by Shepp and Vardi in 1982 [28]. This works on the idea that the statistics of the physical process can be incorporated into the objective function for iterative reconstruction, similar to (2.12), in which assumed the statistical variation observed in the sinogram data was assumed to be Gaussian. However, MLEM works with the principle that nuclear decay is a typical example of a Poisson process, which describes the distribution shown by the number of occurrences of a 'rare' event, $x \in \mathbb{N}^0$, which has a mean occurrence $\bar{x} \in \mathbb{R}^+$. The Poisson distribution is described by the function

$$P(x|\bar{x}) = \frac{\bar{x}^x \exp(-\bar{x})}{x!}. \tag{2.14}$$

This can be used to find the probability, P, of observing outcome x given the mean $\bar{x}$. However, if x is known, but $\bar{x}$ is not, it can also be used to estimate the *likelihood*, L, that the underlying mean is $\bar{x}$ given that outcome x has been observed. This latter case is the definition used in iterative image reconstruction: we observe data, $\boldsymbol{m}$, and seek the underlying mean distribution of activity, $\boldsymbol{f}$. Note that by convention likelihood is denoted $L(\bar{x}|x) \equiv P(x|\bar{x})$. This is to make it clear that $\bar{x}$ is the free parameter.

In a measured PET sinogram, the number of coincidences detected by any given detector pair is a Poisson variable. However, there are actually I measured projections in total, so a mutual likelihood of observing all I independent outcomes is

$$L\left(\boldsymbol{q}|\boldsymbol{m}\right) \equiv \prod_{i=1}^{I} L\left(q_i|m_i\right) = \prod_{i=1}^{I} \frac{q_i^{m_i} \exp(-q_i)}{m_i!}, \tag{2.15}$$

where each q_i is the Poisson mean of that detector pair. Since it is unknown, $\boldsymbol{q}$ can be estimated using the system model in (2.10) involving the image estimate, $\boldsymbol{f}^{(k)}$. The likelihood function is then used to find the most likely set of values for $\boldsymbol{f}^{(k)}$.

The likelihood function is usually transformed into a form known as the Poisson log-likelihood,

$$\mathcal{L}\left(\boldsymbol{q}|\boldsymbol{m}\right) \equiv \log_{\mathrm{e}} L\left(\boldsymbol{q}|\boldsymbol{m}\right) = \sum_{i=1}^{I} \left[m_i \log_{\mathrm{e}} q_i - q_i - \log_{\mathrm{e}}\left(m_i!\right)\right], \tag{2.16}$$

which is easier to handle due to the reduction of the product into a summation. This does not affect the position of the maximum. The image which maximises this objective function via the model of the mean data, $\boldsymbol{f}^{\mathrm{ML}}$, is known as the maximum-likelihood (ML) solution to (2.16). Converging to a solution for the log-likelihood objective function can be achieved using a number of methods, such as gradient ascent. However, convergence with gradient ascent can be slow, and convergence is not guaranteed.

Expectation Maximisation

To address convergence problems, the log-likelihood objective was combined with the expectation maximisation (EM) method, producing the MLEM algorithm [27, 28]. First EM models the expectation of a set of parameters in their 'complete' form, known as the expectation or 'E' step. For PET, this corresponds to defining a set of hidden parameters known as the complete data, $\{z_{ij}\}$, such that

$$m_i = \sum_{j=1}^{J} z_{ij}. \tag{2.17}$$

These hidden variables are designed such that they track which pixel, j, each event in sinogram bin i was emitted from. Since we know already that $\boldsymbol{q}$ is a model of the mean of $\boldsymbol{m}$, we directly infer that

$$\bar{z}_{ij} = a_{ij} f_j \tag{2.18}$$

is a model of the mean of the complete data for the simple data model in (2.10).

In the second and final step of EM, these parameters are then adjusted to maximise this expectation, known as the 'M' step. By comparison of the complete, unknown data with the measured, known data, it can be shown that

$$f_j^{(k+1)} = \frac{f_j^{(k)}}{\sum_{i=1}^{I} a_{ij}} \sum_{i=1}^{I} a_{ij} \frac{m_i}{\sum_{b=1}^{J} a_{ib} f_b^{(k)}}. \tag{2.19}$$

This is known as the MLEM update equation. This equation is used as the staple image reconstruction algorithm in iterative PET reconstruction, because it is monotonic (i.e. every iteration is guaranteed to improve the estimate).

Improving Rate of Convergence

A known problem with MLEM is that, whilst guaranteed to converge, it can take an unspecified amount of time to do so. Occasionally this is circumvented by stopping reconstruction after relatively few iterations. This is also favourable because it limits

the high-contrast noise characteristic of MLEM reconstructions at later iterations.

Alternatively, other reconstruction methods have been reconstructed to accelerate convergence. The most widely used of these is ordered-subsets EM (OSEM), which found that only projecting a subset of data over sub-iterations within each full image update resulted in much faster convergence [29]. However, it is prone to issues such as limit cycles and not converging to the ML maximisor in general [16].

Other methods that sought to accelerate convergence include space-alternating generalised EM (SAGE) methods, which maximise subsets of the complete data model sequentially [30], and the row-action maximum-likelihood algorithm (RAMLA), a modified version of OSEM that is able to converge to the ML image [31].

Maximum *A Posteriori* Reconstruction

First applied to tomography in 1985 [32], Bayes' theorem relates the conditional probability of observing outcomes A and B according to

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \tag{2.20}$$

where $P(A|B)$ is referred to as the posterior probability, which, by Bayesian interpretation, quantifies the 'degree of belief' having accounted for information B. Similarly, $P(B|A)$ is the likelihood function and $P(A)$ is referred to as the prior: the initial degree of belief in A itself. In image reconstruction, we recognise $P(B|A)$ as the Poisson likelihood before, $L(\boldsymbol{q}|\boldsymbol{m}) = P(\boldsymbol{m}|\boldsymbol{q})$. As such, we can arrive at the log–posterior probability

$$Q_{\mathrm{MAP}}(\boldsymbol{q}|\boldsymbol{m}) = \mathcal{L}(\boldsymbol{q}|\boldsymbol{m}) + \log_{\mathrm{e}} P(\boldsymbol{q}). \tag{2.21}$$

with which $\boldsymbol{f}$ can be optimised via the model of $\boldsymbol{q}$. Note that the denominator term has been dropped since it is not relevant to optimisation of $\boldsymbol{f}$. This method is known as maximum *a posteriori* (MAP) reconstruction [16], and is useful because it provides a framework for inclusion of additional information into the reconstruction via the prior term.

Generally it is possible to optimise Q_{MAP} using EM, although there can be difficulties [16]. Alternatively gradient-based optimisers can be used, although these have the usual setbacks. The EM approach was modified in [33] to facilitate the use of MAP estimation using the one-step-late (OSL) algorithm, although this is not guaranteed to converge to $\boldsymbol{f}^{\text{MAP}}$ [16].

Note that what constitutes prior information is relatively open to interpretation. Typically it is used to reduce noise in the image reconstruction, such as use of the Gibbs prior in [32]. MAP is also one of the many approaches used in the literature to introduce motion estimation into the image reconstruction. These methods will be discussed in chapter 4 . First however, the problems posed by motion in PET will be discussed in the next chapter.

Chapter 3

The Challenges Set by Respiratory Motion in PET

3.1 Introduction

This chapter will discuss the challenges set by motion in medical imaging. The nature of motion, particularly respiratory motion, will be outlined in section 3.2. The consequences on PET-based measurements are described in section 3.2.1. Although PET is the focus of this thesis, its low spatial fidelity means that it is usually coupled with other higher-resolution modalities. Motion-related artefacts in these can have consequences when they are used with PET (e.g. for attenuation correction), and so motion effects in MRI and CT are briefly summarised in sections 3.2.2 and 3.2.3, respectively.

3.2 Respiratory Motion

Subject motion can cause problems in imaging-based clinical research. Unfortunately, subject motion can be unavoidable. Furthermore, its clinical impact can be difficult to assess due to lack of a ground truth. This can be particularly true in PET, where the resolution is low and the SNR sometimes poor.

Types of motion can be separated into external motion, such as limb movement [34] or head rotation [35], or internal motion, which includes respiratory motion of the thoracic organs and rib cage [36], beating of the heart [37], and peristalsis [38]. These can be further classified as non-periodic or periodic. The main focus of this thesis is internal respiratory motion of the chest cavity and thoracic organs.

Respiratory motion can be particularly challenging because it is constant and unavoidable. Restriction is not an option, although it can be somewhat mitigated by breath-holding. Although its approximate periodicity can be useful for motion estimation (for the purposes of correction, discussed in the next chapter), the periodicity is only approximate. This can be further worsened by erratic breathing or disease.

A complex relationship between contraction and relaxation of the diaphragm and expansion of the ribcage via the intercostal muscles can introduce variability from cycle to cycle. Each person also has their own way of breathing [39]. This departure from periodicity can be observed between different subjects, or even the same patient on different days [40]. The lack of periodicity can be a problem for any method of motion compensation which depends upon the periodicity explicitly, such as binning data according to respiratory cycle, or parameterised motion modelling involving average cycle motion (discussed in sections 4.2.1 and 4.4 respectively). The types of variation are described using the following terms:

Intra-cycle variation This is the change in motion within each breathing cycle, also known as the hysteresis effect[1]. For example, drawing a breath can be faster than exhaling, or the internal organs could be at different positions at the same point of the respiratory cycle [41]. This causes the subject's organs to follow different paths during inspiration and expiration, affecting efficacy of motion correction algorithms if only a single trajectory is considered. It has been suggested that intra-cycle variation is due to the changes in volume/pressure ratios during inspiration and expiration [42], and possibly due to the additional energy required to inflate the alveoli during inspiration [43].

[1]This name arises from the different paths any given point traces on a velocity/position phase diagram during inspiration and expiration.

Inter-cycle variation This is the change sometimes seen from cycle-to-cycle, such as a deep breath followed by a shallow breath. Inter-cycle variation represents the reproducibility between different cycles. To a first approximation, the size of the displacement of the organs will roughly scale with the depth of the breathing cycle. However, in reality there were will be extra complications. This was studied by comparing a free-breathing model created at one point to another formed 3 minutes later from the same patient. For example, one cause of this variation is thought to be due to the balance of two breathing mechanisms: abdominal and intercostal. The former, which depends on the diaphragm, is usually dominant. For deeper breaths, intercostal muscle contribution increases and coordination between the mechanisms decreases, reducing predictability [44].

Findings in [41] suggest that deviations from average cycle motion due to intra- and inter-cycle variability could be as large as 5 mm. It was also found that these errors were smaller towards exhale and larger towards inhale parts of the respiratory cycle. This suggests that end-expiration breath-hold might be more reproducible than the more comfortable end-inspiration [41].

Occasionally, breath-holding is used to avoid motion degradation in imaging. This is often at end-inspiration or end-expiration, where the former is usually more tolerable for the subject. It is also worth noting that the position held by the body during breath-hold is generally not along the trajectories of free-breathing, especially during inspiration [41]. This can be a source of systematic error when transforming a breath-hold image to other tidal respiratory positions.

Depending on the modality involved, these additional considerations might need to be taken into account during correction. This is discussed further in chapter 4. More specifically, the sensitivity of different modalities depends on factors such as the spatial resolution, the total acquisition time of a scan, and the time resolution of data acquisition.

3.2.1 Motion Effects in PET

Motion in raw PET emission data presents itself as motion blur. This is primarily due to relatively long scan times, since PET scans typically run for several minutes, but can last up to an hour. Therefore, the blur will appear as a superposition of many breathing cycles, such that the data will primarily exhibit blur from the average motion trajectory.

In practice, the character of the motion degradation can include additional motion-related artefacts, depending on the use of other imaging modalities to correct errors in the emission data. For example, attenuation correction with motion-degraded CT can introduce regions of artificially-altered uptake due to mismatching or misidentification of tissues. This is discussed in more detail in section 3.2.3.

PET is only able to resolve displacements up to half the PET resolution, which for current human whole-body scanners is relatively low [20]. Nevertheless, it has been shown that respiratory motion affects PET imaging in the clinic [45]. However, it is anticipated that advancements PSF modelling will improve spatial resolution and therefore increase the degree of motion degradation in PET images [19].

Whilst faster imaging modalities can utilise breath-holding to limit the amount of motion in an acquisition, PET is typically too long for this to be a practical solution. In fact, there is a general lack of ground truth for respiratory motion in PET at all, which can make it hard to qualify the severity or extent of the problem. This can also make it challenging to assess motion correction techniques.

Nonetheless, there have been attempts to assess the impact of respiratory motion in PET, often involving CT or transmission scans as a benchmark, although such studies assume that the lesions will be visible on CT. In [46], it was reported that mean maximal displacement of lung tumours in 22 subjects was 4.5 ± 5.0 mm, as measured with chest radiography. However, 10 patients exhibited no superior-inferior motion; omitting these yielded a mean maximal displacement of 8.3 ± 3.7 mm, with a range of 3 mm to 22 mm. This finding was verified by [47], who reported that tumours in the upper lung displaced by up to 8.7 mm, but those in the bottom of the lung displaced by up to 24.6 mm. However, other studies have shown that CT-visible

lung lesions only displace by 9 mm [36]. The same study also noted a variation in standardised uptake value (SUV, a measure of peak uptake value within a specified region) of up to 30% [36], depending on the respiratory position at the time of CT acquisition. It was found that respiratory motion can lead to a mean overestimation of 1 cm diameter lesion volumes by 1.3 times its real size [48].

More recently, a paper has tried to employ breath-hold techniques to form a ground truth [49]. This was a clinical study involving 95 patients with pulmonary lesions. Alongside a regular 3 minute PET-CT acquisition, patients were asked to hold their breath for as long as possible (observed range was 30 to 143 s). This study found tumour motion ranged between 0 and 45.3 mm (according to gated free-breathing images), with a respiratory period varying between 5 to 20 s. The breath-hold scans exhibited a recovery of $\mathrm{SUV}_{\mathrm{max}}$ between -14.05% and 223.2%, with an average recovery of 51.4%, although it is unclear how the sensitivity of $\mathrm{SUV}_{\mathrm{max}}$ to noise was taken into account.

The clinical consequences of motion generally can involve poor SUV estimation [38, 36], inaccurate localisation of lesions [45], or disappearance of lesions altogether [50]. These effects can lead to mis-staging of disease, issues with radiotherapy planning, and other clinical/research PET applications [51].

Generally, motion correction is not yet a common practice in PET. Applying motion correction can involve the binning of the data into several sinograms, referred to as gates, instead of just one. These are typically binned according to some predefined rule or synchronised motion signal (these are discussed in section 4.4.2, with more details on gating in section 4.2.1). Note that there will still be some motion within each gate, such that image reconstruction of individual gates will still contain some motion blur, albeit reduced. The term of this residual motion blur is referred to as intra-gate motion, versus the inter-gate motion that occurs between any two PET gates.

Most motion correction algorithms aim to minimise inter-gate motion by correcting each gate to a reference position. Ultimately, it is favourable to minimise intra-gate motion too, which can be done by increasing the number of gates in the acquisition. This is associated with an overall drop in SNR in each individual gate, however. In

the extreme, one might choose to avoid gating entirely and correct the displacement of individual events recorded in the listmode data. These concepts will be discussed in greater detail in the next chapter.

Prospects for Reducing Motion Effects in PET

The noisy quality of PET can make it a poor source of motion information for its own motion correction. This means other sources are typically used to estimate inter-gate motion. These can be purpose-built detectors, such as 3D cameras, which are sometimes used for motion correction in brain PET (for example, [52, 53]). Other medical imaging modalities can be used too, such as CT or MRI. This is discussed in depth in chapter 4.

Since PET-CT is the most prolific dual-modality PET scanner, it might be considered the most convenient choice. Computed tomography has good spatial resolution, but it is relatively poor at differentiating between soft tissues. Even worse, motion estimation requires repeated imaging of the subject. Due to the nature of CT, this can significantly increase the administered radiation dose.

PET-MR might prove to be a valuable ionising radiation free alternative, with the extra advantage of good soft tissue contrast. PET-MR scanners currently exist in two forms: sequential and simultaneous. Sequential scanners are separate but co-registered so that one can transform between the coordinates of each FOV. It is questionable whether motion information acquired at a different time of the PET scan will be useful for motion correction[2]. Indeed, there is some doubt on the utility of sequential PET-MR at all, since difficulties appear to currently outweigh applications.

Simultaneous PET-MR shows much more promise than its sequential cousin. These scanners are able to use MR freely during the PET scan. This simultaneous bimodal view of physiological systems might prove to be useful for various applications. In particular, simultaneous imaging allows (in the extreme case) constant motion

[2] As will be seen in chapters below, the proposed techniques in this thesis might be a viable choice of motion correction in sequential PET-MR despite the two datasets being acquired separately.

tracking of the subject using a 'dynamic scan': many quickly-acquired, relatively low-resolution 3D images. Computed tomography cannot boast this ability without the undesirable increase in radiation dose. Also, the acquisition protocol in MRI can be changed to optimise an image for providing motion information. For example, MR can be used to magnetise a grid pattern into the tissue (known as tagging), which will subsequently deform, providing enhanced motion information [54] (although this can prevent the use of other protocols in the mean time).

However, this dedicated use of the MR scanner purely as a motion-correction device may be impractical, since the MR scanner is likely to be required for other clinical (or research) imaging. Instead, MR can be used to arrange the PET data into bins for the purposes of gating, since smaller images in MR imaging can be acquired at a higher rate (see section 3.2.2, below). For example, an MR 'pencil-beam' navigator is a 1D signal, usually over the right hemi-diaphragm for respiratory motion tracking (e.g., [55]) [56]. Alternatively, the centre of k-space can be used if an MR protocol involves radial acquisition trajectories [56]. The specifics of motion correction will be discussed more technically in chapter 4.

3.2.2 Motion Effects in MR

This thesis assumes that the MR part of PET-MR will be used to generate motion estimates for the purposes of PET motion correction. However, MR is prone to motion artefacts too, and care must be taken to avoid these for reliable correction of the PET. Whilst motion correction of MR images themselves is out of the scope of this thesis, it is worth noting some key points that can lead to motion degradation in MR images and how they can be minimised.

Imaging data in MR are acquired as a set of spatial frequencies of the signal in the field of view. This spatial frequency co-ordinate system is referred to as k-space, and can be reconstructed into an image by applying the Fourier transform. The relationship of a point in k-space and one in real space is inherently non-local: varying one point will affect all other points in the opposing space. Generally, the centre of k-space corresponds to low-resolution structures and contrast within the

image, whereas outer k-space encodes fine structures. Many of the effects of motion on MR imaging can be understood by considering the properties of the relationship between k-space and real space, and how the data are acquired [57, 56].

In practical terms, only a single line can be traced out in k-space during a single data acquisition. This can then be repeated with a repetition time, TR, to fill out the entire k-space plane or volume. The path this line traces in k-space is known as a trajectory, and is commonly just row-filling (a 'Cartesian' trajectory), or can be more complex, such as spirals or radial spokes [57]. Many artefacts occur when the time to fill a k-space with sufficient data [56]. If the motion occurs during this time, it can result in inconsistencies in the spatial correspondence of recorded frequencies in k-space. These can lead to blurring or ghosting in the image [58]. Other artefacts that result directly from the physics of the problem, such as erroneous spin manipulation due to tissue displacement [56].

In modern clinical MR, it is common to use gating or triggering to minimise inconsistencies by acquiring data only when the subject is in a given position, although this can sometimes rely on the repeatability of the motion [56]. Alternatively, prospective motion correction can be used, where the gradients used to encode the spatial frequencies are adjusted on-the-fly to compensate for mismatches, such as in [59].

Anything resulting in faster imaging is generally beneficial [56]. For example, parallel imaging, which uses multiple receiver coils to encode additional spatial information (and therefore can be used to require fewer phase encoding steps) has decreased acquisition times (e.g., [60]). Compressed sensing can also be used, reducing the total number of k-space lines required to form an image [61]. Other methods use trajectories which are less sensitive to motion, such as radial phase encoding, minimising the appearance of artefacts [56]. Radial phase encoding techniques have the advantage that they oversample the centre of k-space, which can be used as a signal for regrouping and gating the k-space data [62]. Finally, motion can be estimated and then corrected in the data by treating it as part of the inverse problem of image reconstruction [63].

3.2.3 Issues with CT and Attenuation Correction

Traditionally, PET scanners were attenuation corrected using transmission scans, which showed motion degradation with a similar character to that observed in emission data (i.e. deforming regions exhibited a continuous blur). Contemporary PET scanners, however, use CT for attenuation correction. These, too, are affected by respiratory motion, but the character of the degradation is different to PET. The form of the artefacts depends on the scan time and the temporal relationship between the data acquisition and the respiratory cycle [64].

Computed tomography is typically fast enough that motion artefacts in CT can be circumvented by using breath-holding techniques, although there is no set standard among practitioners on the position used. As discussed above, end-expiration breath-holds are preferred for repeatability purposes. Additionally it has been shown that CT acquired at end-inspiration was shown to change PET quantification significantly from transmission-based attenuation correction, suggesting that end-expiration is also preferable from an attenuation-correction perspective [65]. Breath-holding is not always feasible for heavily debilitated patients, so tidal free-breathing is sometimes preferred [66].

If acquiring during free breathing, it can be beneficial to minimise breathing degradation by shortening rotation times, gating, or corrected reconstruction, although these can have drawbacks such as increased dose, increased scan time, or lower SNR [67]. However, since motion in PET data manifests as a blur across the whole range of motion, attenuation correction with a single motion-compensated CT image can still introduce artefacts, especially around the diaphragm [45, 50].

Finally, it is worth noting that attenuation correction using a sharp attenuation map (whether from CT or MR) will cause the PET image to appear motion-corrected, although this is artificial and care must be taken when using contrast recovery to assess motion correction performance near feature boundaries (e.g. figure 8.1).

Chapter 8, which focusses on incorporating attenuation correction into a motion-estimation algorithm, will discuss this further. For an example of these attenuation effects using simulated data, please see figure 8.1.

3.2.4 A Note on Attenuation Correction and Motion

Whilst no CT is used for the work in this thesis, attenuation correction in PET-MR remains an open question and so assumptions on what form it will take must be minimal. Two assumptions are made with regard to this problem. Firstly, we assume that it is possible to acquire a μ-map (a volumetric image displaying attenuation coefficient values at 511 keV, expressed in cm^{-1}). As discussed in section 2.1.4, some headway has been made in solving the PET-MR attenuation correction problem. A number of current approaches involve acquisition of an MR image which is segmented into tissues (such as water, fat, and air) to produce a μ-map. These will typically be acquired in one respiratory position via a respiratory- and cardiac-gated acquisition because the MR sequence required to image cortical bone (an ultra-short echo time, or UTE, sequence) has a long acquisition time of 11–16 minutes [68].

The second assumption is that the acquired μ-map would be in an end-expiration non-breath-held position *only*. This assumption decouples motion and attenuation into two separate considerations; acquiring an attenuation map in different positions is simply a mix of these two effects. As mentioned above, current MR techniques for attenuation estimation tend to have long acquisition times; acquiring one for many positions would be impractical, and it would make more sense to acquire high-quality MR-based motion estimates and warp a single attenuation map [68].

The discussion in this subsection will become particularly relevant in the final methodology chapter of this thesis, chapter 8. First, however, methods of motion correction must be introduced in the next chapter, alongside a review of respiratory motion correction techniques used in PET so far.

Chapter 4

Estimating and Correcting for Motion

4.1 Introduction

As described in chapter 3, motion can seriously affect the quality of PET images. Motion manifests as continuous motion blur in reconstructed images, and, depending on the attenuation correction method used, can also present as regions of erroneous tracer uptake. In a small number of cases, lesions can become undetectable, although due to lack of a reliable ground truth, the full extent of this is unknown. Recent work by [20] involving simulated data suggests that the problem can be severe for small faint lesions, depending on the subject's breathing.

At the time of writing, there is no widely-adopted method of motion correction in the clinic, although there are several basic methods to reduce motion effects. These methods, known as motion-prevention methods, will be discussed in section 4.2.

In the literature, a host of methods involving estimation of and correcting for motion have been suggested. Inferring motion information from images in these methods is discussed in section 4.3.3. A class of approaches involving modelling subject motion are discussed in section 4.4. Finally, in section 4.5.1, methods of using motion estimates to correct PET images are discussed.

4.2 Reducing Motion Effects

Techniques which minimise effects of motion on the data are amongst the most simple and straightforward methods to implement. Occasionally immobilisation is used, particularly in brain studies [69]. Immobilisation can also be useful for bulk motion, but is ineffective for respiratory motion. Breath-holding in PET has been shown to be possible, although it can be impractical since many PET scans are minutes in length [70, 49].

Gating techniques are possibly the most often used. In cases involving respiratory motion, gating takes advantage of the (approximate) periodicity of the subject's breathing cycle so that data acquisition can occur at the same respiratory position in each cycle. Generally gating can be applied retrospectively, including cases of bulk motion if it is clear that a shift occurred at a given time and was relatively static afterwards. More information on gating is described below.

4.2.1 Gating PET Data

Correcting or minimising the impact of respiratory motion in PET data often relies upon grouping together events recorded whilst the body is in a certain position. Originally referred to as the multiple acquisition frame (MAF) technique [69], the technique eventually came to be known as dynamic framing or gating. The literature can be quite ambiguous on which is meant. For clarity, this thesis refers to PET frames as a series of sequential subsets of data, each corresponding to a unique interval of time. Conversely, gates are triggered according to some rule or signal (generally periodic) and can be considered a summation of multiple frames. This process is illustrated in figure 4.1.

Since PET acquisition is continuous in the time domain, the only thing required to gate the data is a fast and reliable position detection method. Work was done on the optimal number of gates in PET studies, although this was based on abdominal surface tracking [71]. Gating schemes can either derive a respiratory motion signal using an external tracking device or the PET data themselves. Such signals can also

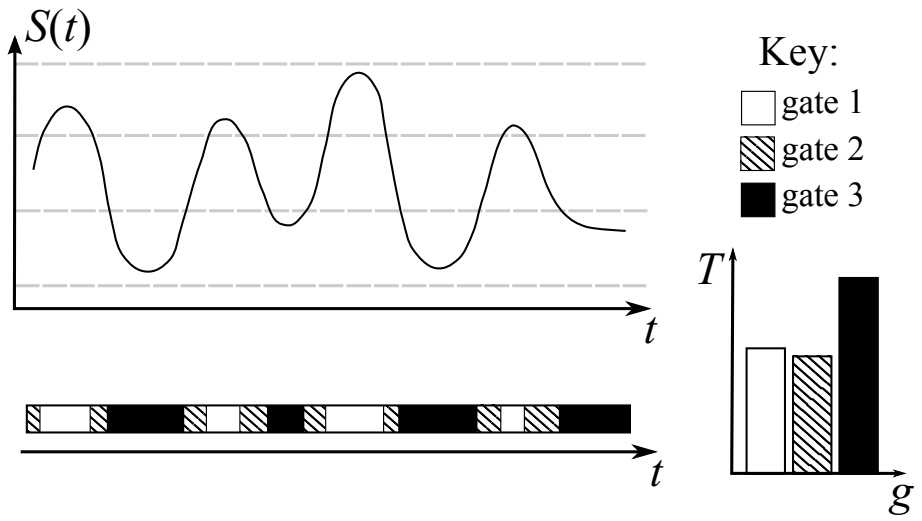

Figure 4.1: **An outline of the concept of PET gating** This figure outlines the idea of gating in PET acquisition. Generally, gating is the separation of data into separate bins, referred to as gates. This is done according to some time-dependent signal, $S(t)$. Conversely, frames involve data acquisition for a single time interval, rather than many. As such, a gate can be considered a superposition of many frames. In the example shown, the signal is used to separate the data into 3 gates. The limits of the gates are indicated by the dashed lines on the graph of S. The corresponding gate that the data is acquired and stored into is shown by the shading of the bar below. The total period of time acquiring into each gate, T, is shown for each gate in the lower right-hand corner. Since gates can vary in total time of acquisition, they can also vary in total number of counts.

be used for parameterised motion modelling, as explained in section 4.4.2 below.

The gating signal can be acquired using external hardware such as respiratory bellows [72], spirometry [73], temperature sensors [74], or video-tracking markers placed upon the patient's thorax or abdomen [75, 51, 71]. However, acquiring a respiratory signal using these methods can be impractical and the correlation between the external signal and the internal motion is not always strong [76]. If the subject is in an MR scanner, an internal respiratory signal can be measured using a pencil-beam navigator, typically positioned on the diaphragm [55, 77]. An advantage of image navigators is that they are very fast to acquire [78] and it was found to be better than other methods of external tracking [79]. However, they require the use of compatible MR protocols, and require continuous use of the MR scanner.

Alternatively, gating can be done using signals derived from the imaging data themselves. Techniques have been proposed to derive respiratory signals from PET using centre-of-mass techniques [80, 81] or spectral analysis of PET sinograms [82, 83, 84]. Machine learning has also been proposed, using principal components analysis (PCA) [85] or Laplacian eigenmaps [86]. A simultaneous PET-MR method has also been proposed, using the intensity of central k-space as a 'self-gating signal' which is used to gate both MR and PET data [62, 56]. Self-gating signals can be a useful means of mitigating the reliance on MR, above, since many protocols can incorporate it. Signals derived using data-derived methods are of interest due to their direct relationship to the internal motion, but they depend heavily upon the quality of the imaging data.

The choice of gating signal and how it is used to bin the data can affect the character of the motion within the gates. For example, binning by amplitude will produce gates which all contain roughly the same range of residual motion, but the number of counts will vary according to time spent in each position (as shown in figure 4.1). Conversely, some methods (e.g. [74]) bin such that the range of events acquired between two amplitude peaks is split into a number of time bins, corresponding to the phase of the waveform [87, 41]. This will ensure the gates will contain the same total counts, but the range of residual motion in each gate will vary [71]. A patient-based cardiac PET study found that phase-like respiratory gating was inferior to

amplitude gating [71].

Whichever method is used, the ideal amount of residual displacement present in each gate should be less than half the intrinsic resolution of the PET scanner. A method for achieving this using a signal and its gradient were recently proposed in [88]. All gating approaches, however, depend upon the assumption of quasi-periodicity of respiratory motion: only an average breathing cycle is obtained in general. Variations in the correspondence between the gating signal(s) and relative positions of internal organs can only be truly addressed using methods which correct individual events (related to intra- and inter-cycle variability of the breathing cycle, discussed in section 3.2).

As a method of minimising the impact of motion on PET images, only one gate (i.e. the one with the most counts) can be selected for image reconstruction, since it will be static relative to the full dataset (e.g. [89, 90]). However, this wastes the data acquired in other gates, and the SNR in the reconstructed image can become very low [74]. To prevent this waste of data, it is more desirable to attempt to motion correct the data instead, allowing use of the full dataset to produce a single, higher SNR image. This requires motion estimates and a method for acquiring them.

4.3 Motion Estimation

Direct measurement of motion is rarely achieved in imaging applications. Instead, it is usually inferred from images. This is done using registration techniques, which seek the transformation, $\mathcal{T}$, that best approximates the transformation from image A, the 'target', to image B, the 'source'. In many cases this will involve formation of an objective function comparing the transformation of target, A', to the source, similar to those used in PET image reconstruction (section 2.2.3). The exact method of inferral depends on the technique in question, and upon the nature of the motion being estimated. How to best represent motion, and then how to estimate it, will be discussed in the following subsections.

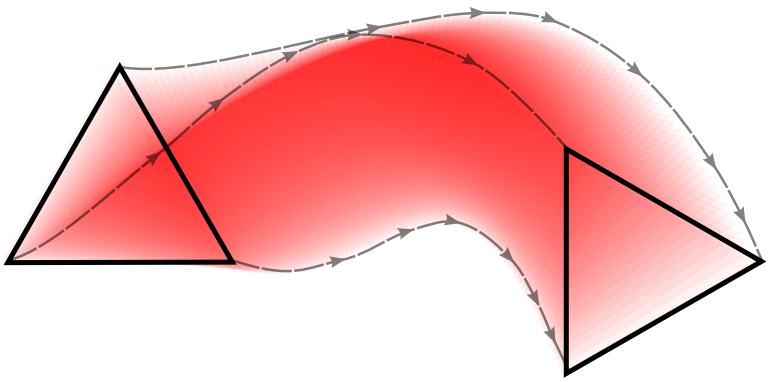

Figure 4.2: **An example of rigid motion** Required complexity of motion estimation depends upon the nature of the motion and the properties of the imaging modality. In this simple example, a triangle of uniform intensity translates left to right, rotating by 90°. The trajectory of each vertex is represented by a line with arrows. The motion blur is shown by the intensity variation: regions containing the triangle for the longest period of time are the most intense. A modality with a long acquisition time but high temporal sampling relative to the motion time, such as PET, would observe this intensity distribution instead of a distinct triangle. Instantaneous acquisitions would only see the triangle before and after motion. This latter case would only require estimation of the direct transformation between the two states, instead of estimating the full motion trajectory, which would be more challenging.

4.3.1 Mathematical Representations of Motion for Digital Images

Consider an assembly of discrete particles in configuration A. As time progresses, each particle will follow a path, until a later time, at which point the assembly will be in configuration B. In this sense, the displacement of each particle is parameterised by time: at any time point, there will be an intermediate configuration of the particles. An example of 3 points, forming the vertices of a triangle, is shown in figure 4.2. Physically, this is what we mean when we discuss 'motion' in a technical sense. Often this is in the limit where there is a continuum of matter rather than a discrete number of particles. In this continuum limit, the motion can be considered as a vector field: each and every infinitesimal point of the object will, generally speaking, have its own motion parameters over time. In figure 4.2, this would correspond to defining a trajectory for each and every possible point on the triangle.

To be able to correct imaging data for motion, an estimate of the motion is required. The actual paths taken by objects may or may not matter: this depends on the frequency of the motion corresponding to the speed and times of data acquisition. As shown in figure 4.2, continuous motion can lead to motion blur which depends on the parameterisation of the trajectories. However, modalities which can acquire quickly and over short periods will only observe snapshots of the motion. In such cases, a motion estimate might be as simple as the linear displacement of each point in the image.

Additional complexity can arise from the nature of the moving object. For example, describing motion with only a translation and rotation will not be satisfactory for objects which change size during the motion. Even more complex are situations involving compression, twisting (relative to surroundings), and sliding along boundaries. The lungs sliding against the chest wall is a good example of this [91]. Such motion is often referred to as nonlinear motion, since it can only be represented by nonlinear transformations. Respiratory motion, especially when considering the transformation of the entire thorax, is a form of nonlinear motion.

As discussed above, the motion of continuous media require a continuous description itself. As noted in the discussion of imaging in section 2.2.2, digital storage of continuous information is not straightforward. Instead the motion must be approximated, with varying success depending on the points discussed above. The following section will outline the main concepts of mathematical representations of motion.

Rigid and Affine Transformations

The simplest motion transformations are 'rigid'. These transformations preserve the relative positions of all points of the transforming object (such as a rotation of the head). This is limited to rotations and translations, or their combination. As a result, the size and shape of the object is preserved. Three-dimensional rigid transformations are mathematically convenient because they only require a maximum of 6 variables to parameterise the transformation of an object. However, in many medical imaging situations the rigid motion description is not a good enough

approximation of the true motion.

Rigid transformations are a subset of affine transformations. Affine transformations are not constrained in the same way; instead they are constrained to preserve parallel lines. This means that they can also represent global changes in scale or shearing of a collection of points. This can be done using homogeneous coordinates, $(x, y, z, 1)^{\mathrm{T}}$, such that

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & \Delta x \\ R_{21} & R_{22} & R_{23} & \Delta y \\ R_{31} & R_{32} & R_{33} & \Delta z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{4.1}$$

where R is the 3D rotation matrix and $(\Delta x, \Delta y, \Delta z)$ are the translational shifts. Affine transformations still only require relatively few parameters to represent a global transformation in an image. For this reason, they are well suited to simple models of motion involving expansion and contraction of individual organs, such as the heart, although not for configurations of multiple organs.

Motion Fields

Affine transformations are limited to global transformations in images. The opposite extreme is the representation of a transformation of every point, locally. These transformations are referred to as non-rigid transformations, which are particularly important for any motion involving local deformations, such as the lungs [92, 93]. Non-rigid transformations can be represented by the deformation or motion field, denoted $\Delta\vec{r}$.

In non-rigid transformations, the deforming object is imagined in such a way that each infinitesimal point has a small arrow associated with it, pointing to where it is going (or, where it has come from, depending on the formulation). This is a form of vector field, akin to that used in electromagnetism. Mathematically, this is represented as

$$\vec{r} \rightarrow \vec{r}' = \vec{r} + \Delta\vec{r}, \tag{4.2}$$

where

$$\Delta\vec{r} = (\Delta x, \Delta y, \Delta z)^{\mathrm{T}}. \tag{4.3}$$

Note that, in this case, $\Delta\vec{r}$ is a function of $\vec{r}$, since it has a value at each and every point. Intuitively, this can be considered as a general spatial transformation between an object at $\vec{r}$ and $\vec{r}\,'$. The application of a motion estimate, $\Delta\vec{r}$, to an image, $\boldsymbol{f}$, according to some predefined motion transformation algorithm, $\mathcal{T}$, is denoted by

$$\boldsymbol{f}_{\mathrm{trans}} = \mathcal{T}\{\boldsymbol{f}, \Delta\vec{r}\}\,. \tag{4.4}$$

where $\boldsymbol{f}_{\mathrm{trans}}$ is the transformed image. The nature of the transformation operator, $\mathcal{T}$, depends on the form of the motion estimates and interpolation of the image.

Naturally, we are limited to a discrete number of points in a digital representation, and motion fields are often stored or represented as a set of N greyscale images for motion in $\mathbb{R}^N$. Each pixel intensity then represents the displacement of that particular pixel in that direction. This follows the image representation formulation discussed in chapter 2: the displacement of each point is supported continuously on a regular grid. Note that a pixelwise representation isn't necessary: many studies reduce the parameters to be stored or estimated by using B-splines, for example [94, 95, 96, 97]. As noted in [98], the performance of B-spline methods depends upon the coarseness of the control point grid. If the grid resolution is much lower than the image resolution, it is not generally possible to model local transformations.

Motion fields can then be used to transform image A into image B by expressing the position of each pixel in 'world coordinates': with a defined origin and expressed in millimetres. Note that this requires selection of an interpolation scheme, which corresponds to re-sampling the grid upon which the image is supported. Moreover, it is worth noting that – due to the limitations of gridded support – motion fields are not directly invertible. Registration can be formulated in a specific way that allows inversion by finding a transformation that optimises some criterion for both forwards and backwards transformations (e.g. using a demons approach [99]), but this can be computationally expensive [100]. Otherwise, the inverse must be estimated using iterative methods [100].

Generalised Motion Matrices

Motion fields are useful for formulation for non-rigid transformations, but sometimes a linear representation of the transformation is required [63]. This is done by representing the transformation of each pixel as a large, sparse matrix, denoted $\boldsymbol{M}$. The intensities in image A are represented as a data $[J \times 1]$ vector, upon which the matrix acts to rearrange the vector elements, corresponding to transformation of individual pixel intensities. This transformed vector is then rearranged back into an image to give the transformed image, A'.

An example of a motion matrix operating on a 4 pixel image, $[\boldsymbol{a}]_j$, is shown in (4.5).

$$\begin{bmatrix} a_1 \\ a_3 \\ a_2 \\ a_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \tag{4.5}$$

The generalised motion matrix representation does not require a list of pixel world coordinates to act on an image, but does still implicitly involve these when calculating which elements of the matrix are nonzero. The example shown in (4.5) uses nearest-neighbour interpolation and therefore consists of only 1s and 0s. Other interpolation schemes will generally produce a set of weights, generally between 0 and 1 [63].

4.3.2 Rigid Motion Estimation

Some of the earliest methods of motion correction in PET involved the estimation of rigid motion. This is particularly applicable to the head [69, 52]. Affine registration was also introduced for motion correction of the heart [80]. In such cases, external measurements are often used to track the motion, such as surface markers with an infrared camera such as the Polaris [101, 102] or feature tracking [53]. In [103], rigid motion parameters were used to correct for respiratory motion of the heart. The motion was estimated from clinical data by minimising the mutual information

between two reconstructed PET gates, followed by rebinning the listmode data prior to reconstruction. This method was then extended to affine transformations [104]. Some of the earlier simultaneous PET-MR motion correction papers also involved rigid transformations for neurological PET, either to rebin LORs [105] or as part of a joint estimation framework [106]. However, rigid motion is not always suitable. More complex deformations of the body require estimation of many more parameters.

4.3.3 Non-Rigid Motion Estimation

As discussed above, earlier techniques sought to find global rigid or affine transformations only – a relatively simple estimation task due to the small number of parameters. More complex tasks, such as respiratory motion, require estimation of non-rigid transformations, which can be much more computationally demanding.

A method of general non-rigid motion parameter estimation in PET was presented in [107]. Other methods vary in the definition of the objective function used to register the images. For example, [108, 109, 110] modelled the deformations as elastic, favouring those which stretch the least. These are similar to other physics-based methods which employ biomechanical penalty terms [111, 37, 112].

Alternative approaches include the use of the demons algorithm [113], Bayesian methods which incorporate a temporal image-based prior [114, 115, 116], and optical flow algorithms in general [108, 115, 117, 118]. B-splines can also be used to smooth the motion fields and reduce the number of parameters to be estimated [94, 95, 97]. For more information on non-rigid motion estimation techniques, please refer to the review paper by Sotiras *et al* [119].

In many applications it is necessary to register a series of images which vary smoothly according to a higher-dimensional deformation model. An example of such a case is registration of respiratory gated PET images, in which the motion between each subsequent gate typically follows a smooth trajectory representing a typical respiratory cycle. When registering such images it can be advantageous to take account of the higher order properties of the deformation. One approach is to use the smoothness as a constraint or regularisation on a series of 3D registrations, each of which estimates

the motion between an image in the series and a reference image (e.g., [80], who applied this to PET gates). An alternative is to employ a full 4D registration, which allows registration of all images simultaneously using a single motion representation that describes the deformations between all images in the series (e.g. [120]).

Typically, inferral of non-rigid motion through registration requires images with a high fidelity of spatial information: high-resolution image features, high-contrast boundaries, high SNR, fine pixel sizes, and so on. However, high noise levels are common in PET. Moreover, since gate-based motion correction methods can only correct inter-gate motion effects, it is usually favourable to push towards more gates to reduce the amount of intra-gate motion within each gate. This can make direct PET-PET registration even more challenging because it decreases the SNR in each gate [117, 95], although steps can be taken to mitigate this.

For example, it is common to use B-spline–based motion fields. These use a set of control points to undersample the motion, reducing the number of parameters required to describe the non-rigid displacement of all pixels. A pixelwise motion estimate can then be generated by interpolating between the control points. The reduced number of parameters in B-spline–type approaches can work better than full voxelwise motion estimates due to robustness to noise [121]. There are associated problems with this, however. In particular, B-splines enforce smoothness and therefore cannot account for discontinuous boundaries in the motion transformation, such as slipping along organ boundaries [122]. Due to the reduced sampling of the motion by control points and limitations of the spline interpolation, they are also poor at modelling localised deformations [123]. Another method used to improve robustness to noise is hierarchical registration, where convergence is stabilised by registering sub-sampled images and steadily increasing the sampling rate for finer registration [98].

Due to the limitations in deriving motion from PET, a modality capable of high fidelity spatial information, such as 4D CT, is often used to motion correct PET data. However, since this involves additional exposure to ionising radiation, CT can only be used to track motion for a limited time. Some works that have used CT to derive motion estimates for PET include [67, 124, 5, 113, 125].

More recently, the introduction of simultaneous PET-MR scanners has opened up the possibility of using MR imaging to estimate subject motion during PET acquisition without the limitations imposed by radiation exposure. Past methods of estimating motion with MR images and applying the results to PET include tagging [126], joint motion and MR-image estimation during reconstruction [127], reordered-2D slices [55], fast 3D volume acquisitions [128], or manifold alignment [129]. A paper was published involving general motion estimation during PET reconstruction, which could also be used with MR, although it is not MR-specific [107].

Work involving these methods on an actual PET-MR scanner is beginning to be published. For example, in [55] respiratory-gated MR images were registered to estimate motion transformations. These transformations were then matched to corresponding PET gates and used to transform them to a reference position. In [62], a respiratory signal was derived from MR and used to form gated MR images for correcting PET gates. In [129], gated MR images were formed from 2D slices using manifold alignment, and subsequently used for motion correction of simulated PET gates. In [130], a technique was described that required less MR scanner time. A respiratory signal was derived from the PET data and used to gate the listmode data into sinograms. The same signal was used to gate the MR data during a short, extra scan after the main PET acquisition. Motion was estimated from the MR gates and applied to motion correct the PET gates.

4.4 Parameterised Motion Modelling

The motion estimation methods discussed above can be generalised by parameterising the motion description. This allows the motion representation to describe a range of motions – rather than just one at a time – and at different positions. These representations are known as parameterised motion models. Typically this is done using motion surrogates (see section 4.4.2) to reduce the dimensionality of the description, although how these surrogates are used to produce an estimate varies between direct- and indirect-correspondence types, described below. A comprehensive review of parameterised motion modelling is given in [4].

Any motion model is only as good as the correspondence between the full- and reduced-dimensionality representations of the motion to the ground truth. Parameterised modelling is particularly useful for respiratory motion, since it is approximately periodic.

4.4.1 Definition

In the literature, the phrase 'motion model' is used for a variety of concepts: sometimes it refers to a motion estimate (which, in some sense, is a model of the motion in the image), and sometimes to the explicit modelling of the overall range of motion transformations as a function of some surrogate signal.

In this thesis the term 'motion model' will only be used in to refer to algorithms which produce fully-descriptive motion estimates using a surrogate, and are therefore able to interpolate between observed motion states. The distinction between direct- and indirect-correspondence models and 4D registration algorithms is explained in section 4.4.5.

4.4.2 Reduced Parameterisation: Motion Surrogates

In some instances involving motion correction, such as PET gating, a full description of the motion is not required. Instead, one only requires a simpler signal correlated to the motion, reducing the number of degrees of freedom to describe the transformation. A clear example is the affine transformation, in the above section: the motion of each and every pixel can be approximated by using a simpler description of global rotations and scaling – this only requires 12 numbers. This concept is also closely related to signals used for gating PET data, discussed in section 4.2.1, where amplitude or phase are used to bin listmode events according to respiratory position.

In some cases, one can explicitly model the relationship between the simpler signal and the complex nature of the overall motion, as in parameterised motion modelling. In this context, the simpler signal is referred to as the motion 'surrogate'. A

surrogate, in this context, is any signal related to the motion being modelled – this can be a scalar signal (such as an amplitude) or data with higher complexity, such as multiple signals, 2D surfaces, or 3D image volumes [4].

For respiratory motion, surrogate signals can be derived using external measurement devices, such as spirometry [73], which measures airflow when breathing, or a respiratory bellows [72], which derives a signal using a bag of air strapped to a patient's chest. 3D visual and infrared cameras (and even a Microsoft Kinect [131]) have also been used. One study found that surface motion estimates are more appropriate for bulk motion estimation since the correspondence between surface deformation and internal motion may be less clear compared to the alternative methods [73]. Another study found that the use of surfaces to drive motion models produced better results than scalar signals [132].

Modalities with good spatial information, such as MR or CT, can also be used to extract surrogate data. This work focusses on PET-MR, making the MR scanner the obvious candidate. Various MR surrogates have been suggested and implemented, requiring varying degrees of complexity. Methods of particular note include 1D image signals ('navigators') [78, 133], 2D surfaces [134], and respiratory phase [41]. Other methods are comprehensively summarised and reviewed in [4]. An example of a navigator signal can be seen in figure 4.1. The processes by which surrogates can be used to estimate overall motion with modelling is described below.

4.4.3 Direct-Correspondence Motion Models

The most intuitive method for creating full-FOV motion estimates from a simpler motion surrogate is one which finds a direct link, or 'correspondence', between the two. Generally, a direct-correspondence (DC) motion model can take the form of a function which takes the surrogate as input and uses it to directly generate the corresponding motion estimate. This could involve, for example, using a scalar motion surrogate to interpolate between a set of measured motion fields [130].

Direct correspondence models can be described as mathematical functions, $\vec{\phi}$, such

that

$$\Delta\vec{r} = \vec{\phi}(\boldsymbol{n}), \tag{4.6}$$

where $\Delta\vec{r}$ is the motion field estimate, and $\boldsymbol{n}$ would be a measured motion surrogate, such as those given as examples in section 4.4.2. The application of the direct correspondence motion model can therefore be expressed as

$$\boldsymbol{f}_{\text{trans}} = \mathcal{T}\left\{\boldsymbol{f}, \vec{\phi}(\boldsymbol{n})\right\}. \tag{4.7}$$

where $\boldsymbol{f}_{\text{trans}}$ is the transformed version of the original image. $\boldsymbol{f}$. Often the literature describes the generation of the DC motion model itself as the *formation* step. The use of the model to produce motion estimates, which are subsequently used to transform an image, is known as the *application* step.

Direct-correspondence motion models can be as simple or complex as required, depending on the situation. A simple example of a DC model for producing an estimate of a motion field in direction k is[1]

$$\boldsymbol{\phi}_k(\boldsymbol{n}) = \boldsymbol{A}_k\boldsymbol{n} + \boldsymbol{\phi}_{k,0}. \tag{4.8}$$

If this component of the motion field has a value at each pixel in a J-pixel image, $\boldsymbol{A}_k$ is a $[J \times S]$ matrix representing a linear transformation of the $[S \times 1]$ values in the motion surrogate, $\boldsymbol{n}$, plus a $[J \times 1]$ constant offset, $\boldsymbol{\phi}_{k,0}$. This $\boldsymbol{n}$ can any measured signal of any dimensionality (expressed as a column vector), which is used to directly estimate the motion from the model, such as the CT-derived surfaces used in [132]. To continue the example, in [132] $\boldsymbol{A}$ encodes the information learned by using PCA on transformations from registered 4D CT images. Other examples of linear DC models include [73, 135], with many more presented in the review paper [4].

A more complex DC model could be based on P^{th} order polynomial models. The

[1]Note that, instead of denoting each spatial component of the data vector separately as $\boldsymbol{\phi}_k$, it is possible to instead denote all directions collectively by $\vec{\boldsymbol{\phi}}$, where the overhead arrow corresponds to the spatial characteristic of the vector. This will be important to minimise ambiguity and improve readability later in this thesis.

k^{th} spatial component of the model is formulated as

$$\boldsymbol{\phi}_k(\boldsymbol{n}) = \sum_{p=0}^{P} \boldsymbol{A}_{pk} \boldsymbol{n}^p, \tag{4.9}$$

where the p^{th} power of $\boldsymbol{n}$ is an element-wise operation, up to P. Polynomial motion models can be more accurate than linear models, although this depends on the relationship between the surrogate signal and the overall motion. In respiratory motion applications, polynomial models usually only set $P = 2$ or $P = 3$ to minimise the risk of over-fitting for higher-order polynomials, which could otherwise lead to large interpolation or extrapolation errors [136]. Some examples of polynomial models include [136] and [1].

The models above can be improved to cope with intra-cycle respiratory motion by having different behaviour for inspiration and expiration. This can be done by, for example, finding a separate set of coefficients, $\boldsymbol{A}$, for both inspiration and expiration. However, it is desirable that the motion estimates produced by such models are constrained to meet (i.e. produce the same motion estimates) at the end-inhale and end-exhale surrogate values. This approach was employed in [137] for a polynomial affine respiratory motion model. An alternative approach was presented in [120]. This model avoids the need for separate models for the inhale and exhale pathways by using respiratory phase instead of amplitude (as is typically the case). Although the use of respiratory phase allows some intra-cycle variation to be captured, this comes at the cost of reduced ability to capture inter-cycle variation because the respiratory amplitude is effectively normalised.

4.4.4 Indirect-Correspondence Models

An indirect-correspondence (IC) model is formed in the same way as a DC motion model, but the assumption that the surrogate directly drives the production of a motion estimate is relaxed. More specifically, one assumes that there are internal variables within the model which can be optimised. This allows the motion model to be more adaptive in its application: surrogates other than those used to form the model can be used to optimise the model according to its internal variables.

In many cases, IC motion models are image-driven: the surrogate data used to optimise the model is a (possibly partial FOV or low SNR) image of the organ(s) of interest. Typically, given some similarity measure, Sim, the internal variables, $\boldsymbol{s}$, are optimised by transforming a reference image, $\boldsymbol{f}_{\text{ref}}$, using a trial motion transformation created by an estimate of the internal variables. The transformed reference image (possibly subsequently processed in some way) is compared to the surrogate image, and their similarity is used to improve the estimate of these internal variables. This can be expressed mathematically by

$$\hat{\boldsymbol{s}} = \arg\max_{\boldsymbol{s}} \text{Sim}\left(\mathcal{T}\left\{\boldsymbol{f}, \vec{\phi}\left(\boldsymbol{s}\right)\right\}, \boldsymbol{f}_{\text{ref}}\right), \tag{4.10}$$

where $\hat{\boldsymbol{s}}$ are the optimised values of $\boldsymbol{s}$, which are used to transform the image according to

$$\boldsymbol{f}_{\text{trans}} = \mathcal{T}\left\{\boldsymbol{f}, \vec{\phi}\left(\hat{\boldsymbol{s}}\right)\right\}. \tag{4.11}$$

In principle, non–image-driven methods could be used as an alternative to the above if some comparison were devised to enable optimisation of the internal variables. Examples of IC motion models include [138] and [139], with other examples provided in the motion model review paper, [4]. In the case of [139], the motion model is first formed by correlating 3D dynamic MR volumes with the motion transformations obtained from registering them. In this case, if the volumes were also used to drive the model, this would be a DC model. Instead, to apply the model, 2D MR slices were acquired and registered to the 3D volumes using the motion model, where an optimal match implied the best motion fields. Therefore, the data used to drive the model (i.e. the surrogate data) is the 2D MR slice, rather than the data used to form the model (the image volumes).

It is the IC method that is predominantly used in this thesis, the specific details of which will be outlined in each methodology chapter. Note that image-driven IC models can be considered as constrained registration algorithms, where the motion data used to train the model in its formation defines the set of allowed registrations.

4.4.5 Motion Models and 4D Registration

Parameterised motion models are conceptually similar to the 4D registration algorithms discussed in section 4.3.3. Both involve estimation and parameterisation of a 4D transformation. However, in 4D registration, the aim is purely to estimate, or model, the motion between the series of images, whereas parameterised motion models aim to use this model to estimate a new motion field based on some measured surrogate data. The measured surrogate data can be the direct input to the model (as in DC models), or it can be used to optimise a (possibly different) signal used to form the model (as is the case in IC models). In this sense, the IC modelling has stronger links to 4D registration algorithms, and it is possible to use the result of a 4D registration to apply a motion model using an IC approach.

4.4.6 Using Parameterised Motion Models in PET

Direct correspondence motion models have previously been proposed for motion correction in PET [140, 1, 130, 127]. There has also been one example of an indirect correspondence motion model for PET motion correction, which used 2D MR images as surrogate data [139]. Although the results were promising, this approach would also require the use of a simultaneous MR scanner throughout PET acquisition, which would be inconvenient for many clinical MR imaging protocols.

To the authors' knowledge, there has not yet been any use of an IC motion model that has used the PET data themselves as the surrogate. If such an approach could be demonstrated to be effective, it would have a significant advantage in that the MR scanner would only need to be used to acquire a short calibration scan for motion model formation at the beginning of the simultaneous PET-MR scanning session. After this, the MR scanner would be free for clinical use. The approach would also have the added advantage of being compatible with sequential PET-MR. The application of an IC model in this way results in one of the key novelties of this thesis: It offers a theoretical basis for enabling the use of both MR and PET data when estimating motion fields for PET motion correction.

In this thesis, such techniques will be described and demonstrated, which use an MR-derived motion model to constrain the range of possible PET-PET registrations, thus addressing the difficulties associated with estimating motion from low-SNR PET gates. Note that the PET data would still need to be gated. This can be achieved using the MR scanner, but there are alternative ways (some of which are deviceless), as discussed in section 4.2.1. There have been very few combined MR+PET-based motion estimation approaches in the literature (e.g. see [106, 141, 97]). These will be discussed in more detail in chapter 7.

4.5 Applying Motion Correction to PET

4.5.1 Using Motion Estimates to Correct PET Data

Given some motion estimate, $\Delta\vec{r}$, there are a number of methods that can be used to motion correct PET images. Generally speaking, each can be categorised into different stages of the PET image formation process: pre-reconstruction, during reconstruction, and post-reconstruction. There were also attempts to introduce motion correction into analytical techniques of image reconstruction (e.g. [64, 142]) although these are beyond the scope of this work.

Reconstruct-Transform-Average

The reconstruct-transform-add (RTA) method is a post-reconstruction process that is the most straightforward to implement. Introduced in [69], RTA requires separate reconstruction of G individual gates, denoted $\boldsymbol{f}_g$, each of which is transformed to the position of the one selected as the reference gate, and subsequently summed to produce a single, motion-corrected image:

$$\boldsymbol{f}_{\text{RTA}} = \sum_{g=1}^{G} \mathcal{T}\{\boldsymbol{f}_g, \Delta\vec{r}_g\} \tag{4.12}$$

There are no restrictions on the form of the motion estimate, $\Delta\vec{r}_g$.

A large number of techniques in the literature are RTA-based due to the lack of complications that arise due to the other incorporation methods. In particular, RTA has been used in several proof of principle PET-MR studies [143, 139, 144, 55].

RTA can also be useful if motion is being estimated directly from the PET because it produces the intermediate, uncorrected images. A number of papers have tried this [108, 75, 109, 110, 141]. In practical settings this method of motion estimation can give rise to complications involving attenuation correction, sometimes requiring a second reconstruction.

Another complication that can arise from RTA is that it requires the reconstruction of individual gates (see section 4.2), which are low in counts and high in noise, and thus suffers from all the associated problems. Reconstruction of many PET gates in this way is unreliable due to the increased noise in the data [5, 145]. In some cases, the final image can also suffer from issues with bias in the motion-corrected image, resulting from the non-negativity constraint in PET reconstruction algorithms such as MLEM [146]. This can, however, be reduced by appropriate regularisation techniques [147, 148].

Motion-Corrected Image Reconstruction

To minimise issues with low counts in the data sinogram, motion correction can be incorporated into the reconstruction algorithm itself, so that all the acquired data can be used to reconstruct the image. A version of this was introduced in 2004 by Rahmim *et al* [149], which produced a single motion corrected image. In 2006, 3 papers published a formal framework involving the incorporation of a number of gates [124, 5, 113].

Although this method has been referred to by a number of names, here it shall be referred to as motion-compensated image reconstruction (MCIR). The idea works on modification of the model of the mean data in (2.10) by factorising the system matrix into a separate geometric term and a transformation term, representing the motion. This can be denoted by the general motion matrices, $\boldsymbol{M}_g$, where g labels which of the G gates the transformation corresponds to. This representation of

motion was described in section 4.3.1. This can then be incorporated into iterative reconstruction framework. For example, the MLEM image update equation becomes

$$\boldsymbol{f}^{(k+1)} = \frac{\boldsymbol{f}^{(k)}}{\sum_g \boldsymbol{M}_g^{\mathrm{T}} \boldsymbol{A}^{\mathrm{T}} \mathbf{1}} \sum_{g=1}^{G} \boldsymbol{M}_g^{\mathrm{T}} \boldsymbol{A}^{\mathrm{T}} \frac{\boldsymbol{m}_g}{\boldsymbol{A} \boldsymbol{M}_g \boldsymbol{f}^{(k)}} \tag{4.13}$$

given the (gated) measured data, $\boldsymbol{m}_g$, and the system matrix, $\boldsymbol{A}$. This has also been done with other reconstruction schemes, such as MAP [150].

There have been several studies into whether RTA or MCIR is better in terms of image quality and quantification. In 2006, [151] found that MCIR exhibits lower variance than RTA for MLEM reconstruction, and it similarly outperforms RTA in quadratic MAP. This was built upon by [145], who used PET simulations derived from real MR data to show that MCIR is better at recovering the correct values in an image due to the bias arising from low count data reconstruction in RTA. However, in this paper it was found that MCIR becomes noisier than RTA with large numbers of iterations. A recent quantitative comparison of RTA and MCIR was published using real PET-MR data [152], which found that MCIR generally performed better in terms of lesion quantification.

A number of papers have been published involving MCIR (for example, [126, 144, 130, 127]). However, a drawback of this method is that it requires the motion estimates to be known prior to reconstruction. There has been some headway into joint motion and image estimation techniques grounded in MCIR-type reconstruction. Groundwork for general motion estimation was developed in by [107]. Several MAP-type approaches have been developed, which include Bayesian-like objective functions [111, 114, 153, 116].

As discussed, all gating-based techniques are fundamentally limited in the amount of motion correction they can achieve; most gates will contain some amount of intra-gate motion. To address this, LOR-based motion correction might ultimately prove to be the best method for motion correction.

Line-of-Response–Based Motion Correction

To bypass the limitations of gated motion correction, it is possible to motion correct the data in projection space. There are two predominant methods of doing this in the literature, and even then the lines between the two methods are blurred.

The first of these is LOR rebinning. The overall aim is to assess which LOR a given event would have been detected to, had no motion taken place. Since LORs are straight lines, this seems like a natural approach for global rigid transformations. Rigid motion–related LOR rebinning algorithms have been employed by [102, 154, 101, 102, 35, 149]. An example of an affine correction technique was presented in [80]. It was found that in a case involving rigid motion correction of the brain, the LOR rebinning technique produced better results [35]. This was also attempted recently for non-rigid transformations [155]. Other methods have also been tried, such as deconvolution [156] and recomputation of the system matrix [157].

Line-of-response rebinning has inherent complications due to the nature of PET data: LORs are straight lines and detectors are only located in some places. This can result in loss of data, should some events be shifted into an unphysical location. Similarly, LORs that were never detected can be brought into the FOV of the scanner. This can lead to image artefacts due to inconsistent information provided by different views. This can be addressed by careful consideration of missing LORs and the changes in sensitivity and normalisation [101, 158, 159]. Rahmim and colleagues sought to address this problem by modelling additional virtual LORs into the reconstruction [149].

The second class of results are those that transform the system matrix in data space, rather than image space as in MCIR. For example, [154] involved no rebinning. Instead, a time-varying data model was constructed and incorporated into the EM algorithm. A similar technique has been employed recently by representing the problem with tetrahedral meshes [160]. Other methods include similar LOR-transformation approaches for either rigid transformations [105] or non-rigid transformations [94, 96, 152]. In particular, [152] found this class of LOR-MCIR–type reconstruction performed better than RTA in a situation correcting respiratory

motion using B-spline–based registrations of MR images.

4.6 Summary

In this chapter, methods of characterisation, estimation, and correction of motion have been described. The work in this thesis aims to complement the motion literature by exploring the viability of reduced-parameter motion estimation by employing MR-derived motion models. Due to the lack of ground truth in respiratory motion problems, it is commonplace to use simulation data to fully characterise the efficacy of new techniques. The next chapter will outline the simulation procedures used for the work in this thesis. The following chapters will then outline how motion models can be used with RTA-type reconstructions to make estimation of motion fields robust to image noise. Later chapters will then focus on incorporating motion models into the reconstruction itself.

Chapter 5

Simulation Framework

5.1 Introduction

As outlined in chapter 2, iterative image reconstruction is based upon a mathematical model of the PET data acquisition process. It is particularly important in PET that the data acquisition model is accurate and reliable to ensure that PET remains a quantitative imaging modality. A typical statistical model, such as that used for MLEM, relates the measured data, $\boldsymbol{m}$, to their mean outcome, $\boldsymbol{q}$, according to the system matrix $\boldsymbol{A}$ and an image of the spatial activity distribution, f, according to

$$\boldsymbol{q} = \boldsymbol{A}\boldsymbol{f}. \tag{5.1}$$

In many cases, the system matrix can be factorised to model additional affects that might affect the acquisition of true coincidence events, such as attenuation, $\boldsymbol{X}$[1], and motion, $\boldsymbol{M}$ [161, 162, 163]. This would give rise to a more complicated system model,

$$\boldsymbol{q} = \boldsymbol{X}\boldsymbol{A}\boldsymbol{M}\boldsymbol{f} + \boldsymbol{\sigma} + \boldsymbol{\rho}, \tag{5.2}$$

where $\boldsymbol{\sigma}$ and $\boldsymbol{\rho}$ are sinograms modelling scatter effects and random coincidences, respectively.

[1] Although attenuation can be represented as a matrix, it can be constructed as $\boldsymbol{X} = \mathrm{diag}(\boldsymbol{a})$, where $\boldsymbol{a}$ is an attenuation factor sinogram.

The models used in this way can also be used as a guide for simulating PET data for research purposes. This is useful because simulations allow gradual inclusion sources of error, such as attenuation, and study how it affects the problem being studied. Perhaps most importantly, simulations allow a ground truth for meaningful comparison of techniques. In this thesis, the ground truth is the tracer distribution without any motion present, which will be referred to as 'motionless' data. The motionless simulations created in this work provide characterisation of best achievable performance of all methods presented.

Positron emission tomography data tend to be simulated using either Monte Carlo or 'analytical' methods. The former uses random sampling to create noisy PET data with realistic Poisson statistics according to the physics of the problem. Geant4 Application for Emission Tomography, or GATE, is a widely-used Monte Carlo PET data simulator [164]. However, Monte Carlo methods tend to be slow, with the time taken dependent upon the required accuracy of the data and the quantity produced. This can be addressed by using analytical simulations instead. These use a data acquisition model such as that above to generate the mean data sinogram, $\boldsymbol{q}$, followed by application of a pseudorandom number generator to generate a single noise realisation of the data. This is then used in place of PET data, $\boldsymbol{m}$. Analytical approaches often involve approximation, but they can be faster and more convenient for tasks such as algorithm development. Furthermore, analytical simulations can be designed such that the model of the data is known, and potentially the same as that used for reconstruction. For these reasons, analytical methods were preferred for all simulations in this thesis.

Analytical techniques require a numerical phantom, which simulates the physiological properties of the underlying subject. A common numerical phantom in the literature is the XCAT phantom (and its earlier counterpart, NCAT), which were designed using multi-modal anatomical data [165] (e.g. [109, 166, 97]). These phantoms were designed by segmenting subject organs and simplified their shape using B-splines[2]. The variation in the control points of these B-splines was then used to

[2]B-spline–based phantoms may pose problems due to the issues with B-spline representation of motion noted in section 4.3.3: B-splines have trouble accurately representing local deformations [123] and discontinuous motion, such as sliding at organ boundaries [122]. Therefore, care should be taken for B-spline–based motion estimation on phantom data derived using B-spline motion

generate an anatomy representative of the 50^{th} percentile of the population sample. These phantoms also included models of cardiac and respiratory motion, which can be used to generate dynamic datasets. These are then used by researchers as-is or transformed into sinograms and used with external simulation software.

A method for analytical simulation of PET data was presented in [128], which uses real MR-based images to generate a 3D subject-specific numerical phantom. Motion can also be measured, ensuring the use of realistic motion estimates during simulation. The resulting images were processed using the Software for Tomographic Image Reconstruction (STIR, [2]), where the base images were projected and Poisson noise was simulated. STIR was also used for image reconstruction. Other researchers have used similar segmented-MRI methods [68, 144, 127].

All experiments in this thesis are simulation based, using numerical phantoms I derived by segmenting UTE MRI volumes of volunteers. The volunteer MRI data were acquired by myself and Christoph Kolbitsch together. The work in chapter 6 followed on from the subject-specific phantom work of [128, 68] with PET simulation and reconstructions done using STIR. STIR simulations were run using code written in Perl, written by myself but with assistance from Harry Tsoumpas and Irene Polycarpou. Later, the decision was made to move away from STIR to avoid the additional complications to algorithm development introduced by learning details of a large, open-source programming project. Therefore, later in the project, simulations were generated using MATLAB because it was more flexible for algorithm development. The MATLAB code was written by myself. These MATLAB-based simulations were initially limited to multiple-slice 2D using the built-in Radon function, but were adapted to 3D using some MEX-based projectors provided by Andrew Reader. The following sections will explain the simulation procedures and their development, with some repetition of the process outlined in [128] for completeness. The specifics of each experiment will be explained in each methodology chapter.

estimates itself.

5.2 Data Acquisition and Preparation

All of the simulation techniques used in this project require a specific set of data to work. These will be detailed below. These simulations use *real* anatomical and respiratory motion information, derived from volunteer MR scans. Thus the main factors simulated are the PET acquisition process and its associated data qualities (such as noise character), which is heavily dependent on the data model used, and the algorithm used to implement it. The overall aim is to start with an image of the true activity distribution, $\boldsymbol{f}$, and apply successive operations in (5.2) until a noiseless version of the data sinogram is generated. The simulated data can then be used to assess the efficacy with which a correction algorithm can reconstruct $\boldsymbol{f}$. As mentioned, the exact data model used for each experiment will be outlined in its respective chapter.

The data required for the analytical simulations are as follows:

1. **A pair of MR respiratory-gated ultrashort echo-time (UTE) volumes:** These are respiratory-gated during acquisition such that the subject is seen in the end-exhale position from free-breathing. The two volumes can be combined to generate a third volume, as described in [68], which makes cortical bone easier to see. All three volumes provide complimentary information for step 4: segmentation.

2. **A number of dynamic 3D MR volumes acquired during free breathing:** Multiple dynamic MR volumes are required with a fast acquisition protocol. Ideally these should be acquired at a rate such that several points are sampled in each respiratory cycle, since this allows simulation of successive cycles as would be observed in real data. My data were gated using a pencil-beam navigator.

3. **Motion fields from registration of the dynamic 3D MR volumes:** The end-expiration dynamic volume in the same respiratory position as the UTE volumes is labelled as the reference (or 'source') volume. The remaining volumes are non-rigidly registered to this source volume. Specific information

on the registration method used in these experiments can be found in section 6.4.3. Note that no motion modelling (according to the definition in section 4.4.1) is used in any simulation in this thesis.

4. **A segmented version of the UTE volumes:** All three UTE volumes are used to create a single segmented volume. In this project, all segmentation was done semi-automatically using ITK-SNAP [167]. Regions were delineated for cortical bone, air, lung, and several types of soft tissue. Lesions can be artificially added at points of interest. This resulted in a general 'PET map', which contained tissues relevant to PET emission and attenuation. Table 5.1 shows the specific tissues segmented for the experiments in this thesis.

5. **Generation of separate emission and attenuation maps[3]:** The PET map was separated into emission and attenuation maps. These were assigned physically- and clinically-relevant values, shown in table 5.1. The emission map is assigned standardised uptake values expected in a typical PET scan. In this project, all emission maps were designed to represent uptake of [^{18}F]-fluorodeoxyglucose (FDG), which is correlated to cellular glucose uptake. The attenuation map is assigned linear attenuation coefficients, referred to as μ-values, which quantify absorption of photons, conventionally expressed in cm^{-1}. For more information about μ-values, please refer to section 2.1.4.

6. **Respiratory-gated emission and attenuation maps:** The motion fields from step 3 were used to transform each map into a number of respiratory positions, depending on the motion effects to be simulated. This is explained in more detail in the following sections.

7. **Blurring the emission map to approximate resolution effects:** This is designed to include resolution-based effects such as non-colinearity, detector size, and positron range prior to annihilation. This thesis assumed a uniform PSF throughout the FOV for simplicity, although in reality this would be spatially-variant.

[3]Note that here the word 'map' is used because these volumes are a noiseless guide used to define different tissues relevant to the simulation process.

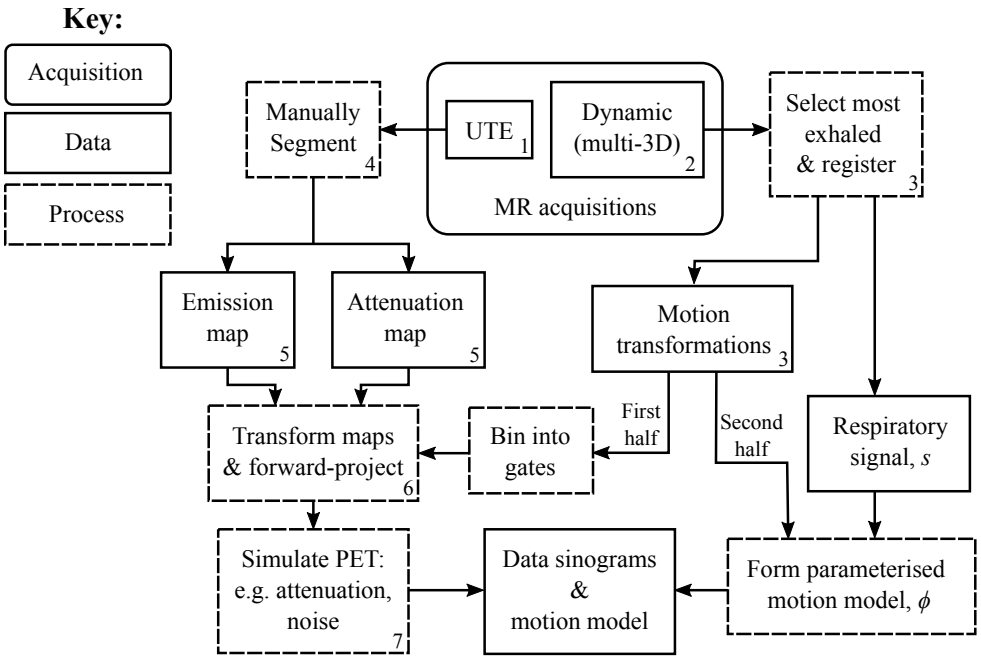

Figure 5.1: **A flow diagram of the simulation procedure** This flow diagram displays the general pipeline used to simulate the PET data used for experiments in this thesis. The numbers correspond to steps outlined in section 5.2 and explained in detail in the text. Note that the motion model formation has been included too for completeness. In summary: UTE images were used for tissue segmentation, particularly bone. These were segmented to generate 3D distribution maps of tracer uptake and linear attenuation coefficient. These maps were transformed into a range of motion states, which were then used to generate noiseless sinograms. Poisson noise was added to the mean data sinograms, proportional to the number of counts expected for each respiratory position.

Please refer to figure 5.1 for a visual representation of how these are derived from each other.

The dynamic MR volumes were assigned to either PET simulation or motion model formation according to order in which they were acquired, but with the constraint that a similar distribution of respiratory states should be present within each group. This ensured that separate data were used for producing the PET test data and for forming the motion model in our experiments, whilst preserving the observed distribution of breathing states, as would be expected in a real acquisition. The method used to simulate intra-gate motion varied with the experiments. In the RTA experiments in chapter 6, the motion fields for each gate were averaged before being applied to the emission maps. In the MCIR experiments, a more realistic approach was used in which a transformed emission map was generated for each

Tissue	Emission Value (SUV)	Attenuation Coefficient (cm^{-1})
Lung	0.05	0.05
Liver	0.2	0.096
Cortical Bone	0.05	0.15
Soft Tissue	0.1	0.096
Myocardium	0.3	0.096
Air	0	0
Lesions	n/a	n/a

Table 5.1: **Simulated values for emission and attenuation properties of various tissues** Values assigned to the emission and attenuation maps, respectively, for each tissue labelled in the segmentation. Note that lesions were also artificially added, but the assigned values varied according to location and experiment. The assigned values are stated in the description of each experiment.

motion field in each bin, then the mean emission map for each bin was used for simulation of that gate. This allowed the PET images to contain the blurred tracer distribution that would be expected in patient data.

5.3 STIR-based Simulations

Simulations for the motion-estimating RTA (ME-RTA) work in chapter 6 were created using STIR, an open-source software written and developed in C++ by Thielemans *et al* [2]. STIR has many tools and utilities for reconstruction, simulation, analysis, and correction of PET data. Notably, there are a number of iterative reconstruction algorithms available, with or without regularisation, along with filtering, motion correction of images, and arithmetic operations can be performed on sinograms. More recently, support was added for single-photon emission computed tomography (SPECT) data too. The software is designed to be run using function calls in the command line.

As outlined above, the main aim of PET simulation is to produce data which represents a set of measured sinograms, since these are a form of raw data a real scanner would output to the user. To do this, each of the terms that contribute towards $\boldsymbol{q}$ in (5.2) must be simulated. STIR has a number of methods for generating the system matrix or projecting on-the-fly, both using a variation of Siddon's algorithm and according to a user-specified scanner geometry. In line with the work by [128], the

scanner geometry used was that of the Philips Gemini TF. The ray-tracing projector was used to forward-project the motion-transformed emission maps, producing a set of noiseless true emission sinograms. This same geometry is used throughout the STIR software, although there is some variation in the algorithm used to project.

STIR can also be used to generate an attenuation factor (AF) sinogram, $\boldsymbol{a}$, from the attenuation maps. Such a sinogram contains the percentage intensity drop for each LOR due to the absorption law (details in section 2.1.4). This provides the $\boldsymbol{X}$ in (5.2).

The $\boldsymbol{\rho}$ and $\boldsymbol{\sigma}$ terms – representing the random background and scatter effects respectively – can also be included using STIR. For example, the former could be included by adding a constant number of counts to each LOR of the trues sinogram with the `STIR_math` utility, which allows basic mathematical operations on the sinograms. The scatter sinogram is more complicated to estimate, and is currently modelled using the single-scatter simulation (SSS) algorithm [15]. This calculation is an analytical calculation derived from emission and attenuation maps, but is still slow to compute, and can be impractical for large numbers of simulations and reconstructions. To first order, this can be approximated using a downsampled image, then interpolating the results back to the original resolution [145]. Note that random and scatter contributions were not simulated in the experiments in this thesis and have only been included here for completeness.

Detection of coincidences in PET follows a Poisson distribution [28]. It is important to accurately represent this noise character in simulated data since it is a defining obstacle in the image reconstruction problem. In STIR, this is achieved using the `poisson_noise` function, which requires a seed number (for the pseudo-random number generation), and a scale factor. The scale factor defines the level of noise, relative to the signal of the noiseless sinogram used as the input. For example, a typical noiseless sinogram modelled in the above fashion might have approximately 1×10^9 mean counts in total, across all LORs. A real, measured sinogram from a 5 minute thoracic FDG scan, on the other hand, typically has approximately total 5×10^7 recorded coincidences [128, 20]. In such a case, the scale factor provided would be 0.05, producing a SNR expected for such a scan but distributed according

to the noiseless distribution. Since the number of counts affect the estimated uptake values in the reconstructed image, the noisy sinograms were divided by the same scale factor. As such, the final noisy sinograms reproduced the expected uptake values in a reconstructed image compared to the original emission maps, but with the noise character of a sinogram with only 5×10^7 counts.

A number of reconstruction algorithms are available in STIR. All experiments using the STIR simulations were reconstructed using the ordered-subsets maximum-a-posteriori one-step-late (OSMAPOSL) algorithm, based upon the MLEM method with optimisations for efficiency. No priors were used in the reconstructions. Note that as an OSEM-based algorithm, this does not generally converge to the ML estimate of the Poisson objective function.

5.4 MATLAB Simulations

The STIR-based simulations were used for creating the ME-RTA datasets, which featured in the oral presentation at the PSMR workshop, the poster presentation at the IEEE Medical Imaging Conference, a paper published in a journal [3], and this thesis. Later in the project, it was decided that it would be less time-consuming and easier to develop reconstruction algorithms using code written specifically for this purpose. This was developed piecewise, with each of the terms in (5.2) written as its own function in MATLAB and, whilst slower than STIR, was more useful for algorithm development.

Initially the projectors used were not based on any particular scanner geometry. Instead, the `radon` and `iradon` functions were used in conjunction with 2D images to produce 2D sinograms since it was efficient, well-tested for bugs, and simple. These functions were easily adapted to produce an approximation of a 3D numerical scanner acquiring only in-plane coincidences, corresponding to segment zero of a Michelogram[4]. This was unrealistic since contemporary PET scanners also have sinograms at angles oblique to the transaxial plane (which form segments beyond

[4]This is also referred to as multi-2D PET, back in the days where lead septa were used to create low-noise 3D images of patients in the clinics

segment zero; for more information on this and Michelograms please refer to section 2.1.3 and figure 2.4). Later, once all the simulation functions were adapted, the forward-projection function was rewritten so that it used projectors of a real 3D scanner geometry. The geometry of this projector matched that of the Siemens mMR. This geometry was selected because it corresponds to the PET-MR scanner recently installed in the PET Centre at St. Thomas' Hospital, allowing easier adaptation to real data in the future. These projectors, given to me by Andrew Reader, were written in C++ and interfaced into MATLAB using MATLAB's MEX toolkit, circumventing some of the limitations introduced by MATLAB's relatively low processing speed.

Attenuation factor sinograms were generated by forward-projecting the motion-transformed μ-maps. Each pixel in the resulting sinogram, $\boldsymbol{\alpha}$, contained the total attenuation coefficient along that LOR. Each AF sinogram, $\boldsymbol{a}$, was generated using this sinogram, according to the attenuation law in (2.3), repeated here for convenience:

$$a_i = \exp\left(-\alpha_i\right).$$

Once all the components had been combined to form the noiseless mean data sinogram (according to (5.2)), noise was simulated. Similar to the STIR function, a noise scale factor was found and multiplied across the noiseless prompt sinograms. Since these formed the mean for the distribution, the sinograms could be inputted directly into the native `poissrnd` function. If the original uptake values were to be preserved, the result of the noise-adding process must be divided by the same scale factor, as before.

Similar to STIR, the same numerical projection functions were used to both create the simulations and to reconstruct the data through whatever reconstruction method required. The projector functions, in the 3D case, were the only non-MATLAB third party software used in the later experiments. The advantage is that development of the reconstruction algorithms could be rapidly modified. The disadvantage is that this methodology is likely more prone to bugs (since it was written by one person rather than a collaboration), and has not been formally tested against real data, unlike STIR [2]. Further advantages and disadvantages are discussed in the next

section.

5.5 Discussion

5.5.1 Comparison between STIR and MATLAB Simulations

The two approaches to simulation outlined above are quite different, and each are more convenient for different situations. STIR is a thoroughly-tested, community-led project, but in its current form it can be quite inflexible without good in-depth knowledge of the classes, functions, and dependencies of the project hierarchy. An advantage to a large project is convenient and fast access to many pre-written utilities and excellent community support, but getting to grips with each function can take some time. The MATLAB code, on the other hand, is less rigorously tested, but for the purposes of algorithm development, is more flexible and faster to implement ideas. MATLAB code can also be slower than the C++ code used for STIR, although the use of MEX files can mitigate that somewhat. Another trade-off when moving to MATLAB was the loss of the level of numerical and geometrical accuracy developed in STIR. However, from a practical standpoint, the leading order (i.e. dominating) effects in the simulations required for this project could be suitably approximated in MATLAB. However, in the future, it would be beneficial to attempt to work the motion estimation functions presented in this thesis into STIR and test them more rigorously.

STIR is implemented using the command line and file storage. This can result in many files being generated across the simulation, which, whilst optimal for computational resources, is not as self-contained as a simulation written in MATLAB. STIR is a useful tool for simulating data reliably and accurately without the extra time required for Monte Carlo methods.

5.5.2 Limitations and Future Improvements

The motion fields used in these simulation methods are voxelwise and derived from real MR scans. The motion should therefore be realistic. However, deriving motion from low resolution dynamic 3D MR scans has known weaknesses. In particular, contrast inside the lungs can be poor. Thus motion fields inside the lungs are, effectively, interpolated from those at the high-contrast lung boundaries. This is likely due to the unreliable motion fields as well as the smaller magnitude of motion in this region. There is some promising work by [129] that might address this difficulty. Increased anatomical detail within the lung provides more information to the registration algorithms, and thus low-information regions within the lung will be smaller than before.

A further limitation of this work involves the timing of acquisition of MR data for simulation. The envisaged use of the techniques introduced in this thesis would be acquisition of the model formation data at the start of, or prior to, the PET scan. Some variation might occur over this period, such that there will be minor differences in the subject's breathing during the PET acquisition. Whilst separate MR volumes were used for motion modelling and for including motion in the simulation, these were acquired during the same dynamic MR scan, and therefore the simulations might not accurately reflect any potential variation in breathing that could be observed in a real scenario. This could be improved by acquiring the data for the motion model and for the simulation at separate time intervals. For example, in this simulation pipeline, the two datasets could be acquired before and after the UTE scan, respectively.

Finally, the work presented this thesis focussed on leading order effects such as noise quality, motion simulation, and attenuation. There are other effects that can be included to improve the quality of the PET simulations, such as improved resolution modelling, scatter simulations, and random background counts. Furthermore, only one realisation was used per PET simulation. This was due to time constraints in processing many datasets with many realisations. However, to assess the statistical behaviour of the algorithms tested in this work, it would be important to repeat the experiments over multiple noise realisations.

Chapter 6

Reduced-Parameter Motion Estimation from Gated PET

6.1 Introduction

As discussed in chapter 3, motion can have a detrimental effect on measured tracer uptake, lesion position, and lesion size. This can result in errors in detectability and quantification, leading to problems in disease staging, radiotherapy planning, and other clinical/research PET applications [51]. There are a number of methods for compensating for motion in data, such as LOR rebinning, incorporating into motion (MCIR), and transforming and summing individually-reconstructed images (RTA). Of these, RTA is one of the easiest to implement. The intermediate reconstruction of PET images can also be useful for techniques which seek to estimate the motion transformations from the images themselves.

However, estimating non-rigid transformations required to correct for respiratory motion can be challenging in PET data alone due to noise [117, 95]. In addition, there must be a compromise between the amount intra-gate motion and the amount of noise in the individual gates. One way to overcome this limitation is to reduce the number of parameters to be estimated from the PET data. This chapter will introduce and assess an RTA-based method which strongly constrains the transformations that can be estimated from the PET images. The research in this chapter

was published in BioMedical Engineering Online [3]. Chapters 7 and 8 describe attempts to incorporate this anatomically-constrained parameter estimation into MCIR, which has a number of advantages over the post-reconstruction method.

6.2 Background

A major challenge in respiratory motion correction is estimating the motion transformations required to correct the images, including the best way to obtain these estimates. The most straight-forward approach is to derive motion information from the PET images themselves. This is particularly the case in RTA, where the generation of multiple PET images, which contain relatively little motion, allows application of a wide range of registration techniques to the images themselves. However, due to the difficulty associated with deriving motion estimates from noisy PET images (for more information, see section 4.3.3), there are relatively few studies on PET-derived RTA methods, many instead tending towards reconstruction-based approaches due to access to all data and access to Bayesian formulations.

One class of non-rigid PET-PET registration involve hierarchical elastic techniques, based on work by [168]. These often involve B-spline parameterisation, with [95, 169] or without [110] regularisation. Elastic registration treats transformations between the source and target image as though they are governed by physical laws of elasticity. This results in a registration objective function resembling Hooke's law, which is equivalent to the use of sum of squared differences (SSD). Care must be taken, however, since SSD will only produce an optimal match when two images only differ by independent and identically-distributed Gaussian noise [170, 168].

An early attempt to estimate motion from PET images in an RTA scheme was performed by [108]. Although this was a cardiac study, it was designed for non-rigid motion estimation in general. This used hierarchical optical flow technique to register two cardiac PET gates. Another class of optical flow algorithms were also tested this purpose [37, 117, 75]. As pointed out by [122], the matching-intensity assumption in optical flow registration might not always apply in non-rigid PET problems since tissue density can vary, and therefore so can the distribution of

radiotracer. This could generate unrealistic motion fields. Mass preservation was built into the optical flow algorithm in [75] to compensate for this [166]. Mass preservation has also been incorporated into a regularised elastic objective function approach [109].

Although registration of PET gates to themselves can be challenging, a study in 2006 successfully registered PET gates using a 4D B-spline registration, which was shown to be robust even with a drop in SNR due to a decrease in acquisition of 12 minutes to 3 minutes [171]. However, the uniform appearance of tracer distributions makes PET unable to estimate twist and stretches, potentially placing a fundamental limit on PET-based motion correction, depending on application [108]. To avoid PET-related motion estimation problems, it could be beneficial to estimate motion from other imaging modalities. For example, this could be done with CT [95]. A significant limitation in CT-based motion estimation is the requirement of additional images, which increases the dose of ionising radiation the subject receives.

To avoid the problems associated with CT, MRI can be used instead. MRI-based RTA techniques include elastic registration methods [127], motion estimation through MR image reconstruction [127], and a Bayesian objective function [141]. Other MR-corrected RTA methods involve multi-2D MR images [55] or 3D radial MRI [62]. There are also methods which involve estimation of motion from PET data but applied during reconstruction, such as [94]. These are discussed in more depth in sections 4.3.3 and 7.1.2.

A notable MR-based motion estimation technique is the use of parameterised motion models [1, 139]. In the case of respiratory motion, these can exploit the approximate periodicity by linking a motion surrogate to the overall more complex motion. In [1], this was applied by directly measuring the surrogate (an MR pencil-beam navigator) and using it to drive the motion model. Using a motion parameter in this way makes this method a direct-correspondence (DC) technique (described in section 4.4.3). Alternatively, in [139], used MR imaging data as the motion surrogate, optimising the internal variables of the model – this is a form of an indirect correspondence model, discussed in section 4.4.4.

The indirect correspondence method in [139] motivated the work in this chapter:

building an MR-based model, but applying PET images as the surrogate. As such, this work compromises between the PET-based and MRI-based classes of motion estimation. MR data is used to constrain the motion positions the PET data can select, but flexibility is allowed in that choice; the only assumption is that each PET gate is in a position *somewhere* along the motion trajectories observed using MR. Additionally, the motion model approach drastically reduces the number of motion parameters that need to be estimated from the PET, reducing the impact of noise on the motion estimation. The rest of this chapter will describe how the feasibility of a PET-based indirect correspondence motion model was examined.

6.3 Method

In this section, the methodology will be described. Specifically, the formation of a respiratory motion model from dynamically-acquired MR volumes will be explained, as well as its application using an indirect correspondence model (see section 4.4.4) to constrain registrations between PET gates. Details of the MR and PET imaging requirements are provided in section 6.3.1. Section 6.3.2 describes how MR images are used to form the motion model. Section 6.3.3 outlines the application of the MR derived motion model using an indirect correspondence model and PET gates as the surrogate data.

Note that throughout this section, the term 'surrogate' refers to the data used to derive a motion estimate from the motion model. Examples of surrogates are given in section 4.4.2, and the general background of using surrogates to produce motion estimates is described in section 4.4. This terminology is used to remain consistent with the motion-modelling literature [4]. In the following experiments and in the context of forming the motion model, the surrogate is a scalar respiratory signal derived from the MR images. When applying the model, two surrogates are used and compared. The benchmark method uses the same scalar signal to directly estimate motion fields. The novel method being introduced uses gated PET data as the surrogate, indirectly estimating the motion fields via an internal variable.

6.3.1 Imaging Requirements

The MR imaging requirements for the motion model used in these experiments are:

- a short dynamic 3D MR scan of the thorax during free breathing, resulting in a temporal sequence of quickly-acquired 3D images (volumes) depicting the thoracic region at arbitrary respiratory motion states;
- a simultaneously-acquired respiratory signal for model formation.

For information on specific acquisitions used in these experiments, refer to section 6.4.1.

The PET imaging requirement is that respiratory-binned PET gates are acquired using a gating technique. This can be done on-the-fly, or retrospectively with a listmode acquisition. This could, for example, be based on an external signal [51] or make use of the PET data itself [82, 172, 86]. The techniques in this chapter are assessed using PET simulations, the details of which are discussed in section 6.4.2.

6.3.2 Motion Model Formation

The formation of the parameterised motion model used in this project is illustrated in figure 6.1, and follows the general overview of motion model formation in section 4.4. The first stage is to estimate respiratory motion using each dynamic MR volume. This is done by applying a non-rigid voxelwise registration [173] of the t^{th} dynamic MR volume to the reference MR volume (corresponding to the most exhaled image, selected using the respiratory signal), which results in a motion field, $\Delta\vec{\boldsymbol{r}}_t$. This vector-like representation is denoted with an arrow above the variable for clarity; this thesis also uses data vectors, which will be denoted $\boldsymbol{v}$ since they have no explicit spatial dependence. A voxelwise motion description was used to avoid the problems associated with B-spline motion modelling mentioned in section 4.3.3 and because the parameter reduction by the motion model prevents any issues arising from PET image noise.

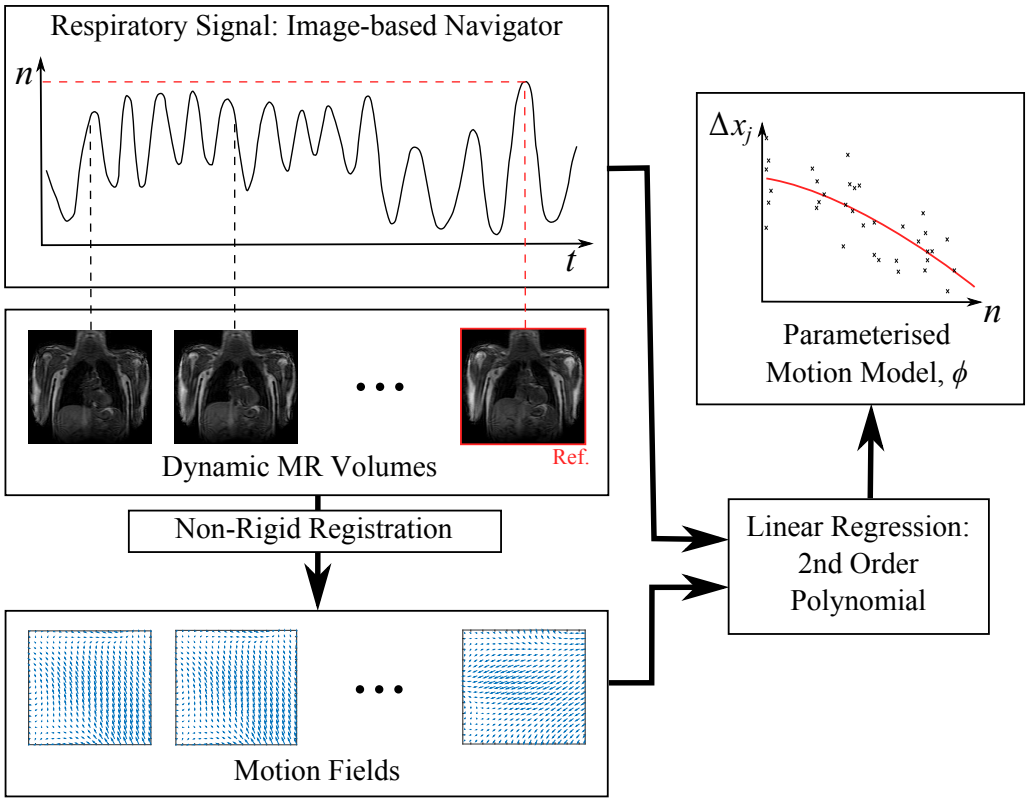

Figure 6.1: **Forming a parameterised motion model** This figure illustrates the method used to form a motion model used in this work. A set of MR volumes is acquired, the most exhaled of which is set as the reference image. The other images were registered to this image to obtain motion fields. A scalar respiratory signal was derived from the MR volumes, which was used with the motion fields to form a polynomial model of the displacement in each direction, for each voxel.

Each dynamic MR volume has an associated value of the respiratory signal, n_t, where $\{n_t\}$ is a set of samples of the continuously-varying respiratory signal n. In this case, n is chosen to be the head-foot displacement of the left hemidiaphragm (see section 6.4.1 for details of how it was measured). These values are used as the independent variable in a regression analysis to form a 2^{nd} order polynomial function of n for each voxel, in each direction. This collection of functions is the motion model, denoted $\vec{\boldsymbol{\phi}}$, which can be expressed in full according to

$$\left[\vec{\boldsymbol{\phi}}(n)\right]_j = [\phi_{jx}(n), \phi_{jy}(n), \phi_{jz}(n)] \tag{6.1}$$

where

$$\phi_{jk}(n) = a_{jk}n^2 + b_{jk}n + c_{jk} \tag{6.2}$$

where indices j and k denote the voxel and direction, respectively. The coefficients a_{jk}, b_{jk}, and c_{jk} are the part of the motion model which encodes the MR information from the formation step. In general there will be $9J$ known coefficients in this J voxel motion model, which are used in conjunction with n to generate voxelwise motion estimates. These polynomial coefficients were determined by linear regression using the Vandermonde matrix method [174].

As discussed in section 4.4, applying this model by acquiring n directly during a simultaneous PET-MR scan would be equivalent to a DC motion model, in which n would be the surrogate. The main objective of the experiments in this chapter, however, is to test whether PET images can be used to estimate the optimal motion fields without acquiring n. This is a type of IC model, for which the PET images are considered the surrogate, since they are the data used to derive motion estimates.

6.3.3 Motion Model Application

To apply the model, the end-expiration PET gate is designated as the reference gate, $\boldsymbol{f}_{\text{ref}}$. This is assumed to be the same respiratory position as the end-expiration MR image used for generating the motion fields for forming the model. This was selected because the end-expiration position is the position that shows the least overall inter-cycle variability [41]. Note that, due to the way that the gates are

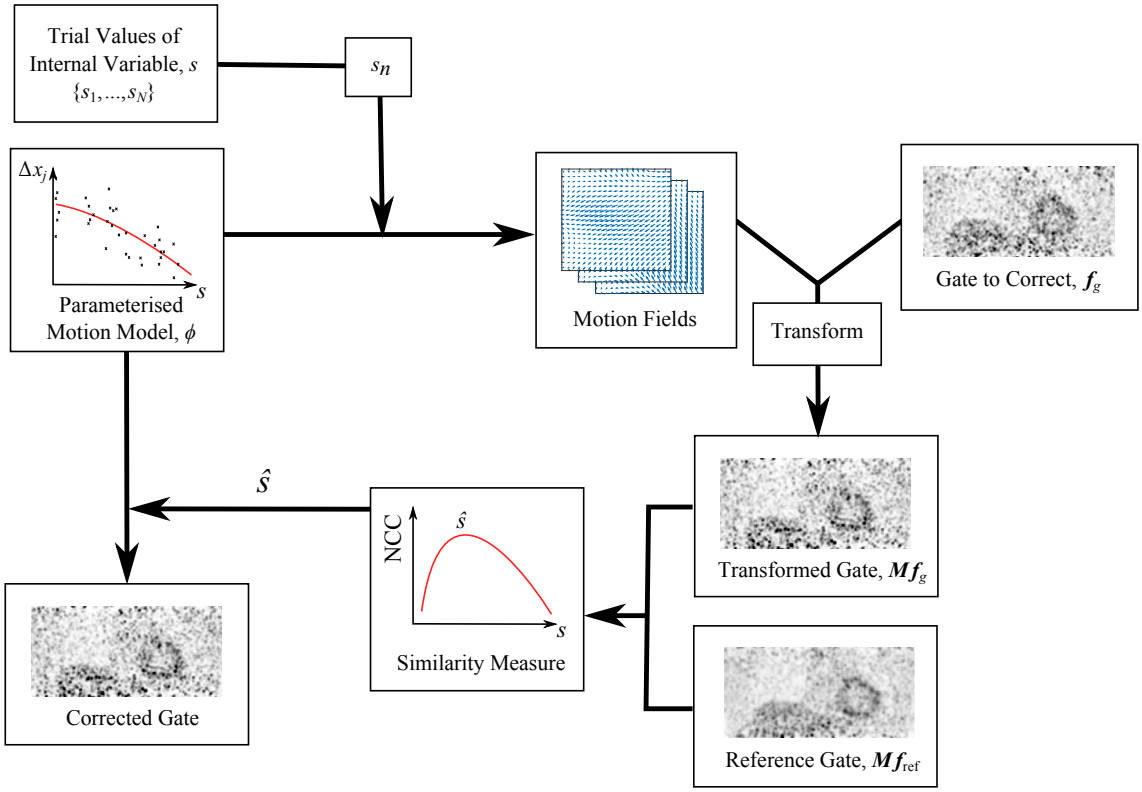

Figure 6.2: **Applying the indirect-correspondence motion model using PET images** This flow chart outlines the algorithm for estimating motion fields using the indirect-correspondence motion model. A number of trial internal variables are used to generate motion estimates with the motion model, which are then used to transform the current gate. Normalised cross-correlation is then used to quantify the success of the trial input. After attempting all trials, the optimal internal variable is used to correct the PET gate.

ordered, $\boldsymbol{f}_{\text{ref}} \equiv \boldsymbol{f}_G$, where G is the number of gates. The g^{th} PET gate to be corrected, $\boldsymbol{f}_g$, is transformed using motion fields produced by the motion model applied with trial values, $\{s_i\}$, of a scalar internal variable, s. This scalar variable corresponds (but might not be identical) to the respiratory signal, n, that was used for model formation, and its trial values are evenly spaced between the minimum and maximum observed values in $\{n_t\}$. Since there is only one internal variable in this implementation, it is not computationally demanding to perform the optimisation using 100 values in an exhaustive search. In more complex implementations, a more sophisticated optimisation could be employed instead. An example of a more sophisticated optimisation approach can be seen in chapter 7, where this approach is incorporated into PET reconstruction.

Each resulting trial-transformed gate is then compared to the reference PET gate. This comparison is performed in a volume of interest (VOI) placed over the lower right lung and the liver, as indicated in figure 6.3. In principle, the VOI could be any size or shape. This region was selected to maximise the contribution of high-contrast, high-motion regions to the similarity measure. With the general theory outlined in chapter 4, this approach is an exhaustive-search implementation of (4.10), using a scalar, s, as the internal variable $\boldsymbol{s}$, and the PET gate, $\boldsymbol{f}_g$, as the surrogate image. Equation 4.10 then becomes

$$\hat{s} = \arg\max_{s} \text{Sim}\left(\boldsymbol{f}_{\text{ref}}, \mathcal{T}\left\{\boldsymbol{f}_g, \vec{\boldsymbol{\phi}}\,(s)\right\}\right). \tag{6.3}$$

The similarity between the two images within the VOI is quantified by normalised cross-correlation (NCC) [170]. Normalised cross-correlation was chosen because amplitude-gated PET images have varying numbers of counts in each gate, determined by the breathing pattern of the subject. Therefore, it is important to use a similarity measure that is insensitive to changes in image intensity.

Figure 6.2 illustrates how the motion model, $\vec{\boldsymbol{\phi}}$, is used in an indirect correspondence model approach to estimate the motion for each PET gate, $\boldsymbol{f}_g$ (the same procedure is used for the other gates) and to use it for PET motion correction. The internal variable, s, takes the place of the respiratory signal, n, that was used in the model

formation. The range of possible values of s is determined by the observed range of values of n, but the correct value for s is found by optimisation, so the respiratory signal does *not* need to be acquired to apply the model. This fact is an important feature of the technique and has a number of potential advantages and disadvantages, which are discussed in sections 6.5 and 6.5.3 respectively.

All PET gates were blurred using a Gaussian filter prior to comparison. This can be beneficial to avoid local minima in the similarity space since the similarity measure in (6.3) might not be cope well with low SNR. Note that this is only possible because the small-scale motion information is contained within the MR-derived motion model. Standard PET-PET methods would not have this freedom to reduce noise in the PET data without affecting motion estimation accuracy.

The motion field that results in the maximum value of the similarity for each gate is selected as the optimal motion field, found using $\widehat{\Delta\vec{\boldsymbol{r}}}_g = \vec{\boldsymbol{\phi}}\left(\hat{s}_g\right)$. The final motion-corrected PET volume, $\hat{\boldsymbol{f}}_{\mathrm{IC}}$ is formed by applying the estimated motion fields to each of the original, unfiltered PET gates and summing them according to the RTA motion correction method:

$$\hat{\boldsymbol{f}}_{\mathrm{IC}} = \sum_{g=1}^{G} \mathcal{T}\left[\boldsymbol{f}_g, \vec{\boldsymbol{\phi}}\left(\hat{s}_g\right)\right], \tag{6.4}$$

where G is the number of PET gates acquired, including the reference gate.

6.4 Experiments and Results

To evaluate the IC approach, simulated PET data created from real MR data were used. This allowed a quantitative evaluation of the performance of the technique using realistic motion fields. Magnetic resonance data were acquired from four healthy male volunteers (ages 22-33). Details of data acquisition are provided in section 6.4.1. The STIR-based PET simulation procedure is described in section 6.4.2. Section 6.4.4 describes the evaluation of the technique using these data, and qualitative and quantitative results are presented in sections 6.4.5 and 6.4.6 respectively.

6.4.1 MR Data Acquisition

Two different MR sequences were acquired in the same scanning session for each volunteer. An ultra-short echo time (UTE) sequence was used to acquire images for forming emission and attenuation maps for PET simulation (see section 6.4.2). A 3D dynamic sequence was used to acquire images for two purposes. Half of the data were used to form the motion model (see section 6.3.2). The other half were used to transform the emission and attenuation maps into real breathing positions for PET simulation. All data in these experiments were acquired using a Philips Achieva 3T MR scanner.

For the UTE sequence, two images were acquired in an interleaved fashion (each with different echo times) and respiratory-gated in the end-expiration position. The details of this sequence can be found in [68]. The FOV was 400 mm × 400 mm × 400 mm at a resolution of 2 mm × 2 mm × 2 mm, with $TR/TE_1 = 6.5/0.14$ ms and $TE_2 = 4.6$ ms. A flip angle of 10° was used. Gating was achieved using a pencil-beam navigator positioned on the right hemidiaphragm. Scan duration was typically 10 to 30 minutes, depending on the subject's breathing pattern and the efficiency of respiratory gating[1]. The two resulting UTE images were subtracted to create a third image, which shows increased cortical bone contrast [68].

35 dynamic 3D MR volumes were acquired in quick succession for each subject during normal tidal breathing. These were used to estimate motion fields for model formation as described in section 6.3.2 and to transform the PET maps into real breathing positions. The sequence used to acquire the volumes was: T1-weighted FFE using SENSE protocol with SENSE factor 8, flip angle = 10°, FOV 500 mm × 450 mm × 245 mm with acquired image resolution 1.5 mm × 4.1 mm × 5 mm (FH, RL, AP) and reconstructed image resolution 1.5 mm × 1.5 mm × 5 mm and a time resolution of 0.7 s per image. For more information on this protocol, refer to [1].

[1]The UTE protocol used in these experiments automatically set the acceptance range of diaphragm positions for data acquisition. This was occasionally updated to a new position according to a moving average. The amount of time actually spent acquiring data depended heavily on the regularity of the subject's breathing: if it was erratic, the diaphragm would spend a smaller proportion of time in the acceptance region and data would therefore be acquired more slowly. Regular breathers, on the other hand, made efficient acquisition times possible.

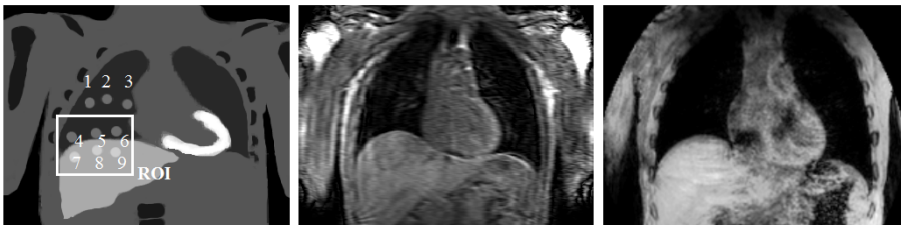

Figure 6.3: **Example of the data used for simulation** This figure shows an example of an emission map, left, with illustrative locations of the 9 lesions for each volunteer. All lesions were placed in this pattern, referred to as lesions 1 to 9 from top left to bottom right. The volume of interest used for motion estimation can also be seen. In the centre is a coronal slice of a dynamic MR image used to generate motion fields from volunteer 3, and a UTE of the same volunteer is shown on the right.

The position of the right hemidiaphragm in the head-foot direction was estimated from each dynamic MR volume using cross-correlation of intensities within a manually-defined, cuboidal VOI [175]. The VOI was comprised of $1 \times 1 \times 20$ voxels over the dome of the liver, where one VOI was compared to a reference VOI by translating in the head-foot direction to find the optimal match. This displacement was then converted into millimetres and assigned as the navigator value for that image volume. The navigator was used as the independent variable in the polynomial regression used for motion model formation. It was subsequently involved in finding the optimal internal variable value during model application, where the PET images themselves were considered to be the surrogate data of the motion model. Refer to sections 6.3.2 and 6.3.3 respectively for more information on model formation and application. The respiratory signal was also used to select a reference end-expiration image for motion estimation.

Note that in these experiments the signal for model formation was image derived, but in practice it could easily be acquired as a pencil-beam navigator by the MR scanner.

6.4.2 Simulating PET Data from Real MR Images

The simulations were based on real anatomical and respiratory motion information derived from volunteer MR images, an approach which has been implemented in

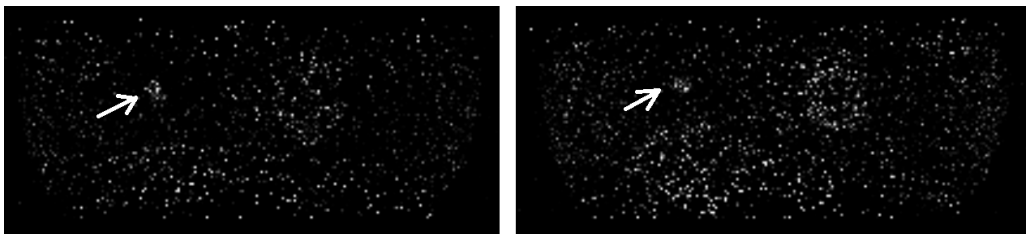

Figure 6.4: **Example gated PET data** Here two amplitude-gated images can be seen, simulated and reconstructed using STIR. The gate on the left is the end-inhale gate, which typically has fewer counts since less time is spent here during the respiratory cycle. On the right is the end-expiration gate, which typically has the most counts for the converse reason. A lesion can be just about seen in these images, indicated by an arrow.

previous studies [128, 68]. The following is an outline of the technique used for work contributing towards the current chapter. For more information on the simulation process, refer to chapter 5.

The UTE image volumes in section 6.4.1 were used to create a manually-segmented map of body tissues relevant to the PET imaging process. This map was further split into a map showing attenuation features (the 'attenuation map') and another showing emission features (the 'emission map'). The emission map was copied 18 times for each volunteer, with one FDG-avid lesion artificially included in a specific location in each: 2 diameters (10 and 14 mm) and 9 locations were used, distributed in the right lung and in the liver. These positions are shown in figure 6.3. These maps were then transformed into different breathing positions using the motion fields derived from half (17/35) of the set of dynamic 3D MR image volumes. In the simulations, 6 respiratory gates were defined. To see how the 17 motion fields were distributed amongst the 6 PET gates (to simulate intra-gate motion), please refer to chapter 5.2. Note that there is no motion modelling (according to the definition in chapter 4) used in this simulation procedure.

The emission and attenuation maps were forward-projected and combined using STIR [2]. This software allowed accurate simulation of PET data and noise effects without the use of time-consuming Monte Carlo methods. 50 million counts were simulated across all gates in each scan. Scatter and random effects were not included. Attenuation maps were assumed to be in the correct motion positions for each gate, and were used to correct for attenuation effects prior to the reconstruction. STIR

was used to iteratively reconstruct each PET gate individually, using the ordered-subsets maximum-a-posteriori one-step-late (OSMAPOSL) algorithm. Note that the MAP part of the algorithm was not used in any reconstruction.

A total of 72 motion-included simulations were created (9 lesion positions, 2 diameters and 4 volunteers). Some examples of the simulated PET gates can be seen in figure 6.4. Note the different noise characteristics in each image due to different count totals in the gated sinograms, explained in the figure caption. A gold-standard 'motionless' (ML) image was also created for each simulation, with the same physical effects and processes. For this, the simulation procedure was identical to the simulations involving motion, except each motionless gate was the reference PET gate used in the motion-included simulations, repeated 6 times (i.e. once for each PET gate). Each ML gate was simulated with the same SNR as the motion-affected equivalent (i.e. each gate contained a number of counts proportional to the time spent in its respective range of breathing amplitude). This was done for consistency, since it accounted for any reconstruction issues resulting from potentially–low-count reconstructions in RTA motion correction. These motionless simulations allow characterisation of best achievable performance of all methods presented in this chapter.

6.4.3 Image Registration

As noted above, the hierarchical adaptive local affine registration ('lreg') algorithm was used to register images in this work [173]. This is a non-rigid registration algorithm, which is necessary to find the free-form transformation required for respiratory motion tissue deformations. These are discussed in general in section 4.3.3.

The lreg algorithm works on the assumption that non-rigid transformations can be split into a series of nested affine transformations. More concretely, a global affine registration is attempted between two image volumes. These volumes are then split into sub-volumes, each of which are locally affine-transformed to the target image. This is continued using successively finer sub-volumes until a stopping condition is met, such as a success measure or the maximum iteration number. The total

non-rigid transformation is then calculated by recursively applying each of these affine transformations to the full image volume. A more technical description of the technique can be found in [173].

This algorithm was used to register and transform the dynamic MR volumes acquired in section 6.4.1 to derive motion fields, which were then used either for PET simulation or motion model formation. It was also used for direct registration of simulated PET gates, which were used as a comparison for the model-constrained registration approach presented in this chapter. In both cases, the default registration parameter values provided on the lreg website were used [176]. Only the similarity measure parameters were varied: NCC was used instead of SSD, with a similarity threshold of 0.999. This was to account for intensity variations between registered images, which was particularly important for the PET gates (as explained in section 6.3.3). This also applies to the experiments in other methodology chapters below.

6.4.4 Evaluation using Simulated PET Data

To characterise the performance of the indirect correspondence model based method (IC) in correcting for the effects of motion on the simulated PET data, it was compared to three alternative approaches to combining multiple PET gates:

- *Uncorrected (UC)*: All 6 motion-affected PET reconstructed gates were summed without including any motion correction.
- *Direct correspondence model (DC)*: The same motion model as that described in section 6.3.2 was employed to motion correct the simulated PET gates, but a DC model–based technique was used for model application. Specifically, the MR image-derived respiratory signal described in section 6.4.1 was used directly as the surrogate input to the model to estimate the motion field for each gate (i.e., in (4.6), the respiratory signal, n, was used as the surrogate). This is the technique described in [1] and would require continuous use of the MR scanner to acquire the pencil-beam navigator. It should form a suitable benchmark for the application of a new motion model.

- *Unconstrained PET-PET Registration (PT)*: Non-rigid registration [173] was used to register each PET gate to the reference PET gate. Details are in section 6.4.3 above. The estimated motion fields were used to motion correct each gate and the transformed gates were subsequently summed via RTA.

Visual inspection was used, plus three methods of quantifying the performance of the technique: peak lesion uptake, lesion size, and position relative to that in the motionless (ML) image.

Peak Lesion Uptake Value

The first quantitative measure was the peak lesion uptake value (specifically, $\mathrm{SUV}_{\mathrm{peak}}$) relative to the corresponding value in the motionless simulations. Peak lesion uptake values of the lesions were calculated by using a small VOI (in this case, a central voxel and its 6 nearest neighbours) and calculating the mean intensity of the small volume. This average was then attributed to the central voxel. The voxel with the highest mean defined in this way within a larger VOI was designated as the $\mathrm{SUV}_{\mathrm{peak}}$ of the lesion. For this work, the large VOI was a user-defined volume of approximately $20\,\mathrm{mm} \times 20\,\mathrm{mm} \times 20\,\mathrm{mm}$ centred on the location of each lesion in the motionless image. Manual definition was required in some cases to avoid noise features within the image. Intensity recovery was quantified as a percentage of the $\mathrm{SUV}_{\mathrm{peak}}$ of each motionless lesion.

Lesion Width

Respiratory motion can cause significant changes in apparent lesion volume and shape. This is seen predominantly in the head-foot (HF) direction since this is the primary direction of displacement for respiratory motion. However, the use of profiles through the lesion does not adequately characterise the shape of the activity distribution. This makes full-width at half-maximum (FWHM) derived from line profiles a poor quantifier of PET motion correction.

Instead, a minimal bounding box (MBB) was defined within the VOI used to find

SUV_{peak} of each lesion. This box minimally fit the whole surface defined by the lesion's FWHM in the 3 directions of the image planes. This was automated using MATLAB, then checked manually. The algorithm created a 3D binary image around each given lesion, in which voxels were assigned a value of 1 if their intensity value was equal to or greater than half the value of the lesion's SUV_{max}. The algorithm then identified the smallest box required to enclose the binary image. The widths of the box were then used to quantify lesion width in each of the 3 image directions: head-foot, anterior-posterior, and left-right. Lesion width recovery was defined as a percentage of the size of the MBB in the motionless simulations, in each direction.

Lesion Position

Finally, the error in lesion position was computed as the third measure. This was measured by comparing the voxel location of SUV_{peak} for a lesion and its motionless simulation, converting to millimetres, and calculating the magnitude displacement.

Note that different Poisson noise realisations were used for each simulation (i.e. the seed for the random number generator was selected randomly between 1 and 1000 for each PET gate), so even if perfect motion correction transformations were applied there would be differences in noise characteristics between any given lesion and its motionless equivalent.

6.4.5 Qualitative Results

Visual Inspection

Some results of applying these three approaches, together with the 'ideal' motionless (ML) simulation are shown in figures 6.5, 6.6 and 6.7. Visually, the proposed approach and all comparative techniques improve the visibility of the lesions. In figure 6.5, which displays a small lesion in position 4, the uncorrected PET image has an indistinct patch of increased activity within the right inferior lobe of the lung. It is questionable whether this would be identified as a lesion by a clinician inspecting

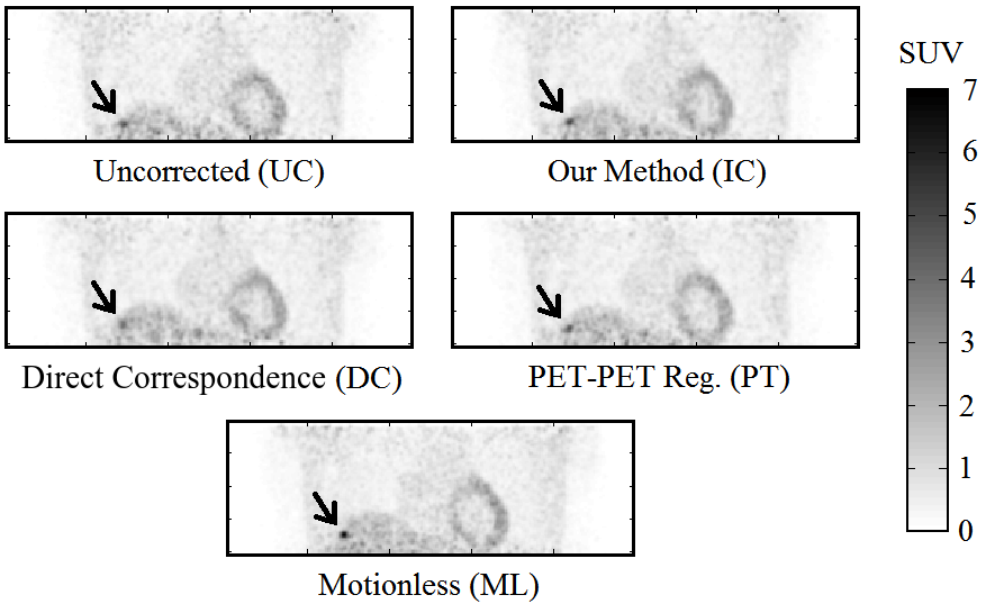

Figure 6.5: **Coronal views of an image containing a small lesion, reconstructed using different methods** This figure displays coronal views of a small lesion, position 4 for volunteer 2, displayed with a **4 mm** Gaussian filter. The intensity scale is shown to the right. The lesion (indicated with an arrow) is relatively indistinct in the UC case, but clear in the IC case. In this example, the PET-PET registration (PT) has performed moderately well, whereas the DC method has performed poorest of the 3 correction methods. Also notice the blurred appearance of the liver-lung boundary in the UC image.

this scan. Upon motion correction, this patch becomes a distinct, higher contrast lesion above the diaphragm. Compared to the motionless case, the motion correction techniques qualitatively recover lesion size and position, but not full contrast.

6.4.6 Quantitative Results

Quantitative results for the experiments are shown in figures 6.8, 6.9 and 6.10. Note that all results are quoted in median and interquartile range since skew was observed in the data distributions. A 2-tailed Wilcoxon signed rank test was applied to test the statistical significance of the results in all experiments. Tests were performed to compare all motion correction techniques to the uncorrected case. A value of $p \leq 0.01$ (i.e. 99% confidence) was taken as a statistically-significant result.

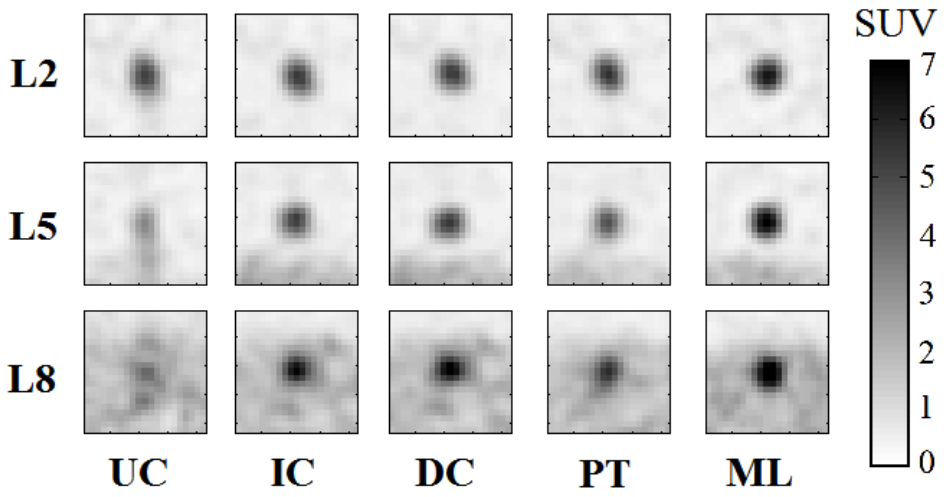

Figure 6.6: **The effects of motion correction on large lesions at various positions for volunteer 4** The effects of motion correction on large lesions in positions **2**, **5**, and **8** in coronal views of volunteer **4**. Please refer to figure 6.3 for lesion positions. Effects are shown for the uncorrected PET (UC), the indirect correspondence model assessed in this chapter (IC), the direct-correspondence method from [1] (DC), unconstrained PET-PET registration (PT) and the gold-standard set by motionless PET (ML). Note that L2 does not move much. This agrees with the observation reported by [47] that lesions in the upper lung move by around 2 mm, which is below PET image resolution. However, L5 and L8 have lost contrast due to respiratory motion, which is recovered with varying success by correction methods IC, DC and PT. Profiles of these lesions are displayed in figure 6.7.

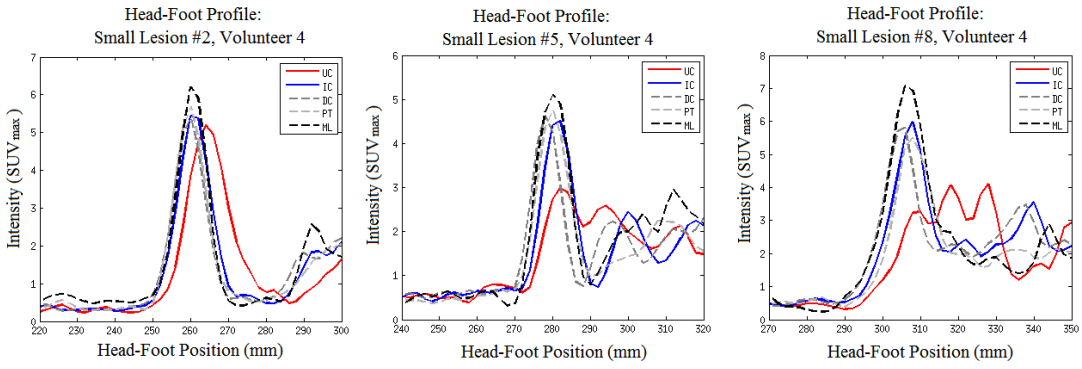

Figure 6.7: **Examples of superior-inferior lesion profiles** Head-foot profiles through large lesions in positions **2**, **5**, and **8** respectively for volunteer **4**. Coronal slices of these lesions are displayed in figure 6.6. Notice the similarity of all profiles for the lung lesion (2, left), but the significant spreading of the UC lesion profile on the diaphragm (5, centre) and in the liver (8, right). In all cases, the IC, DC, and PT correction methods recover a significant proportion of the ML peak uptake value.

Peak Lesion Uptake Value

Recovery of peak uptake value was found using SUV_{peak}. These values are percentages with the motionless lesion at 100%. Values will thus generally be lower than 100%, although sometimes they may be slightly higher due to image noise.

The distribution of changes in SUV_{peak} compared to the motionless case are presented as box and whisker plots in figure 6.8. In the 'All Lesions' graph of figure 6.8, the overall changes in SUV_{peak} of each method are displayed: uncorrected (UC) PET, PET motion corrected with the indirect correspondence technique proposed in this chapter (IC), direct-correspondence motion model (DC), and direct application of non-rigid registration [173] to each PET gate (PT). Subsequent graphs in the figure correspond to each region of the thorax with simulated lesions: the lung (i.e. lesions 1 to 3), diaphragm (lesions 4 to 6) and liver (lesions 7 to 9).

Uncorrected, all lesions were observed to only recover a median of 78.4±18.6% of the motionless lesion SUV_{peak}. The indirect correspondence motion modelling method proposed in this chapter recovered a median of 86.9±13.6% ($p = 1.9 \times 10^{-9}$), whereas the direct-correspondence application of the same motion model yielded 86.3±12.1% ($p = 9.9 \times 10^{-6}$) of the motionless SUV_{peak} peaks. Non-rigid registration of the PET gates recovered $87.2 \pm 16.9\%$ ($p = 3.8 \times 10^{-7}$) of the peak intensities on average.

The lowest observed uncorrected lesion intensity was 40.6% of noiseless activity. The IC, DC, and PT methods recovered this to 78.4%, 64.8%, and 77.3% respectively. The scatter plots in figure 6.12 show the relationship between individual lesion measurements before and after correction. The dashed line $y = x$ defines the point at which no improvement is achieved. Points below this line have worsened under the correction method, whilst those above have improved. These plots are discussed further in section 6.4.7.

Lesion Width

The changes in the head-foot width of the MBB for each lesion FWHM are shown in figure 6.9. Once again, the respective motionless lesion profile was used to define

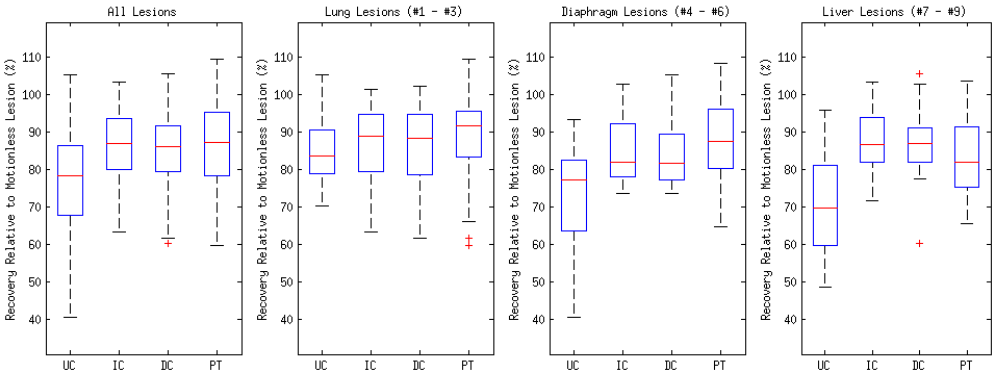

Figure 6.8: **Peak lesion uptake results** Labels correspond to: (UC) Uncorrected PET, (IC) The proposed indirect correspondence technique, (DC) The direct-correspondence technique, and (PT) unconstrained PET-PET registration

100%. In contrast to the results for correction in peak uptake value, the values in figure 6.9 tend to be greater than 100%.

Overall, the uncorrected PET images exhibited a median increase in width to $150 \pm 82\%$ of that of the motionless lesions. The IC motion correction reduced this to $100 \pm 29\%$ of the motionless lesion width ($p = 4.4 \times 10^{-15}$). Similarly, the DC motion correction technique reduced this to $100 \pm 23\%$ ($p = 3.4 \times 10^{-11}$), whereas the direct PET-PET registration reduced the MBB width to $114 \pm 30\%$ ($p = 3.9 \times 10^{-15}$).

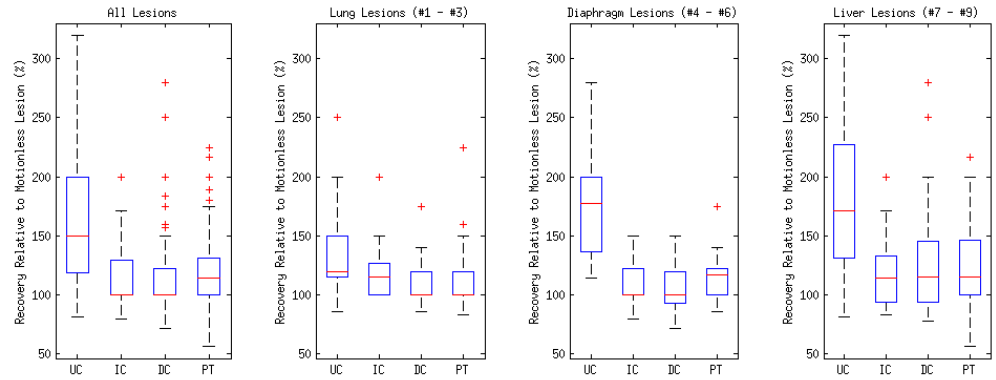

Figure 6.9: **Lesion width results** Note that widths are quoted as a percentage of the corresponding motionless lesion value. Labels correspond to: (UC) Uncorrected PET, (IC) The proposed indirect correspondence technique, (DC) The direct-correspondence technique, and (PT) Direct PET-PET registration. Note that for IC and DC in the All Lesions graph and DC in the Lung Lesions graph, the median corresponds to the lower quartile.

Lesion Position

The displacement magnitudes of lesions from their motionless positions are presented in figure 6.10. Note that the original position of a lesion is at 0 mm. Uncorrected PET lesions showed a median offset of 6.6 ± 5.4 mm. The IC method reduced this median to 3.5 ± 1.8 mm ($p = 2.4 \times 10^{-17}$). The DC method reduced the median displacement to 2.7 ± 2.8 mm ($p = 2.0 \times 10^{-15}$). The PT method reduced the median displacement to 2.7 ± 1.9 mm ($p = 3.6 \times 10^{-18}$).

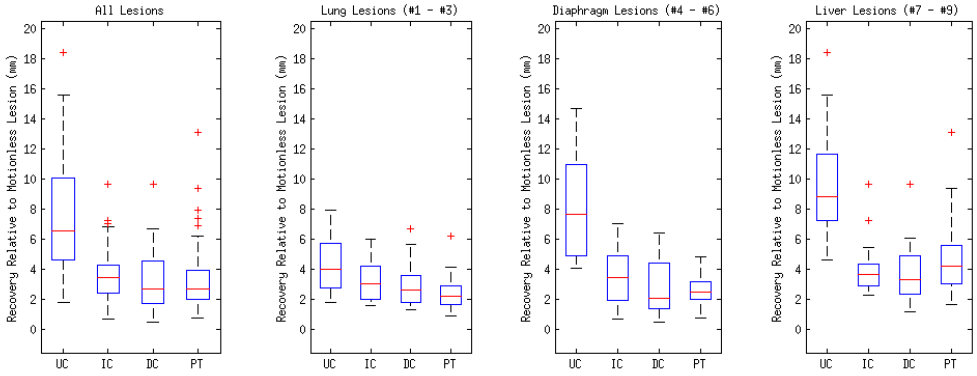

Figure 6.10: **Lesion position results** Labels correspond to: (UC) Uncorrected PET, (IC) The proposed indirect correspondence technique, (DC) The direct-correspondence technique, and (PT) Direct PET-PET registration

6.4.7 Robustness Analysis

The IC technique matched the results achieved by the DC technique, therefore it should be able to achieve correction in any situation for which the DC technique has already been tested, such as those in [1]. Similarly, the IC technique matched unconstrained PET-PET registration, but was more robust in certain cases:

- The range of errors on each PT all-lesion average are larger than those in the IC case, with (occasionally many) outliers on the box plots in figures 6.8, 6.9, and 6.10.

- The PT method underperformed for most liver lesions, and in some cases lesions were not recovered, such as in figure 6.11. There are no cases where a

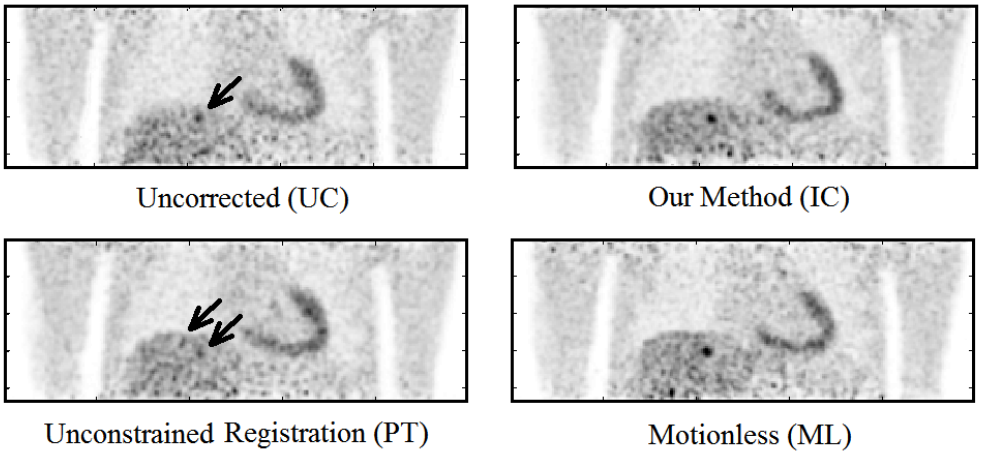

Figure 6.11: **Illustration of poor motion correction of PET-PET registration** Here, the PT method has failed to recover a small lesion in position 9 of volunteer 1. As can be seen, a small, indistinct lesion is indicated by an arrow in the UC image. This is clearly recovered using the proposed method (IC) to a visual quality comparable to the motionless case. However, unconstrained registration (PT) has actively made the lesion harder to see, with two possible candidates indicated by arrows. These are comparable to the levels of noise seen elsewhere in the liver, and could easily be missed. All images are shown on the same intensity scale.

lesion was notably worse due to the IC technique, as can be seen by the scatter plots in figure 6.12. Note that this latter figure only shows lesions in positions 4–9 for clarity; the lung lesions (positions 1–3) are mostly unaffected by the attempts at motion correction due to their small displacement as a result of respiratory motion, and so fluctuate around $y = x$.

- Registration for the IC and DC cases are constrained to a subset of realistic positions, as measured by the MR. The same cannot be said for the PT method, especially when a greater number of respiratory gates are used, due to lower SNR.

6.5 Discussion

A novel RTA-based technique for motion correction of PET gates has been presented. This technique would not require sustained use of the MR scanner during a simultaneous PET-MR scanning session. Performance was comparable with an

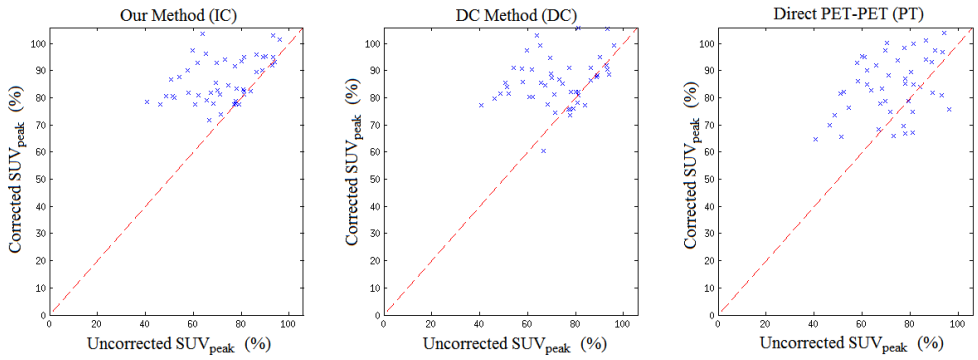

Figure 6.12: **Illustrating method robustness with scatter plots of SUV$_{\text{peak}}$ recovery** These scatter plots show the improved robustness of the two MR-derived motion model based techniques (DC and IC) compared to the direct PET-PET registration technique (PT). Each plot compares distributions of SUV$_{\text{peak}}$ for each lesion in positions 4–9 for all volunteers, both before correction and after correction by the IC, DC and PT methods. The dashed line represents $y = x$. Any point on this line was unaffected by the correction attempt. Points above the line have had their SUVpeak increased (with 100% corresponding to the motionless peak uptake), whereas points below the line show a reduction in SUVpeak. Note in particular that whilst none in the IC case are significantly below the line, 8 lesions dropped in SUVpeak with the PT method (16.7% of the points shown).

MR-derived motion model technique that would require such use of the MR scanner. Experiments indicated that overall the technique could recover median lesion peak uptake value up to $86.9 \pm 13.6\%$ (from $78.4 \pm 18.6\%$ for uncorrected PET), and median lesion size down to $100 \pm 15.7\%$ (from $179 \pm 63.7\%$ in uncorrected PET) of the affected head-foot width. Lesion displacements were improved from 6.6 ± 5.4 mm prior to correction to 3.5 ± 1.8 mm. All of these improvements were statistically significant – below the $p = 0.01$ threshold.

Whilst the average lesion intensity only improved by an additional 8%, this could be due to the depth of breathing of the volunteers. The maximal observed displacements of the right hemidiaphragm in volunteers 1, 2, 3, and 4 were 25.2 mm, 20.7 mm, 13.3 mm, and 38.7 mm respectively. Since the lesion sizes were 10 mm and 14 mm, these displacements span 1 to 3 times the lesion size, which could cause different behaviour under motion. The advantages and disadvantages of this method are discussed below.

6.5.1 Relation to Other Work

The technique described represents a novel approach to PET motion correction, but has some similarities with related approaches from the literature. Several papers have recently proposed the use of MR to obtain motion estimates to correct PET [55, 62]. These works registered MR gates to directly obtain motion estimates, which were then used to motion correct gated PET images prior to averaging them [55, 62]. A similar approach was taken by Manber *et al* [130], in which the respiratory signal used for gating was derived from the PET data. This allowed the MR scanner to be free for clinical use, apart from a short calibration scan (similar to this technique). However, in the work by Manber *et al*, the PET data were not directly used in the motion estimation procedure. In this work, the PET data is directly employed in the motion estimation, but this is constrained using an MR-derived motion model.

4D registration approaches (e.g. [80]) involve a similar 'constrained registration' approach, such that sets of transformations that are not smooth between gates are effectively excluded from the motion estimation process. However, the IC approach uses constraints estimated from another, more reliable, modality (i.e. MRI). The technique does not currently use between-gate smoothness as a constraint, but this would be an interesting extension.

Other related works include [106], who used a probabilistic model based on a mixture of Gaussians to make use of MR data to jointly estimate activity and motion parameters in a single PET reconstruction. More recently, [141] used a combined MR and PET similarity measure when registering gated MR and PET images.

6.5.2 Advantages

The IC technique – based on an IC motion model – has some advantages over other motion correction techniques. The main advantage is that the motion correction only requires the MR scanner for a short, initial motion model calibration scan to provide a method for robust PET-PET registration. After this scan, the motion model can be formed and applied without any further requirement of the MR scanner, maximis-

ing its availability for other clinical or research purposes. This independence from MR also makes this method of respiratory motion correction suitable for sequential PET-MR. Errors in surrogate signal acquisition for DC model–based methods are also avoided. This final point could explain why the IC method occasionally outperformed the DC method in the experiments, resulting in a larger spread of values for lesions corrected by the DC method.

Note that the IC technique does still require a respiratory signal to be measured during MR scanning (for model formation) and also during PET scanning (for gating purposes). However, the two signals do not need to be the same. The requirement to acquire the same or a similar signal, which is a feature of many alternative techniques (such as [140, 1]), introduces several potential difficulties. First, if the signal is measured using the MR scanner, it restricts the use of the technique to simultaneous PET-MR, and furthermore it limits the use of the scanner for clinical purposes during PET scanning. Second, if the signal is measured using an external device such as an optical or magnetic tracker the measurement device must be MR-compatible and any line-of-sight issues must be resolved. Use of an external signal also increases the cost of the solution and complicates clinical workflows. The fact that the IC technique eliminates this requirement is therefore an important feature, and could result in greatly simplified clinical workflows. In addition, using the PET data itself as the surrogate enables the IC technique to use information about PET-visible lesions in the motion correction process, potentially improving motion correction accuracy in areas of clinical interest. However, the biomechanical constraints of motion model based techniques are preserved, leading to more robust results than the direct PET-PET registration technique.

6.5.3 Limitations, Disadvantages, and Future Improvements

The motion fields used in the PET simulations were derived from real MR scans. The motion should therefore be very realistic. However, deriving motion from low resolution dynamic 3D MR scans has known weaknesses: in particular, contrast inside the lungs can be poor. Thus motion fields inside the lungs are, effectively, interpolated from those at the high-contrast lung boundaries. This fact makes the

results for lung lesions (i.e. in positions 1 to 3) less reliable than those close to the diaphragm (positions 4 to 9). Also note that the UC errors in the lungs are smaller than in the other regions. This is likely due to the unreliable motion fields as well as the smaller magnitude of motion in this region.

The IC method uses a VOI to compute the similarity measure from the PET gates. This volume is extended across the lower lung and much of the liver. In the experiments, a lesion in positions 4 through 9 could therefore provide additional information to the IC constrained registration, which could aid the registration process. In contrast, lesions 1 to 3 are outside of this volume of interest. A smaller region was used to maximise the proportion of high-contrast structure within the VOI. This made it easier to estimate the optimal value of the internal variable, $\hat{s}$. Theoretically, the IC registration would find the same motion fields for each of these 3 lesion positions (i.e. 1 to 3) for any given volunteer, since the anatomical information available is identical. However, the noise characteristics of each simulation differ, causing fluctuations in the registration result. Due to this, the IC method performs slightly differently for each lung lesion.

In the IC experiments, perfect attenuation correction was assumed. Attenuation was included in the simulations, but anatomically accurate attenuation maps were assumed to be known for each respiratory position. This allowed the focus of the evaluation to be on the effects of the algorithm on motion correction alone. As it stands, IC would be significantly complicated by attenuation correction using a single attenuation map. Generally, the μ-map would only anatomically match one gate, possibly introducing motion-related attenuation artefacts into the others. This, in turn, could affect the ability to reliably estimate motion during the comparison step since the image features, such as the apparent placement of the diaphragm, will complicate matters. Additionally this method would require twice the number of reconstructions. In chapter 7, the incorporation of this indirect-correspondence technique into image reconstruction via an MCIR-based motion correction algorithm will be introduced. This necessitates formulating the optimisation of an iterative reconstruction to update the internal respiratory signal values for each gate. Incorporation of attenuation correction into the reconstruction procedure is then described

in chapter 8, the inclusion of which makes the assumption of perfect attenuation correction unnecessary.

The motion model employed in this work was a relatively simple 'average-cycle' model, which would not be able to capture any intra-cycle motion variation and only limited inter-cycle variation. Use of this simple model allowed the demonstration of a proof-of-principle for the proposed motion correction approach, but more sophisticated types of motion model could result in improvements in performance. For example, multiple internal variables could be employed, such as signals derived from other anatomical positions like the chest/abdomen or even statistical dimensionality reduction approaches such as principle component analysis [85, 177].

A further limitation of this work is the fact that the MR data used for model formation and PET simulation were acquired in a single acquisition. It would be more realistic to acquire the data in separate acquisitions (perhaps 10 or so minutes apart) to simulate any variation in breathing that might occur in between the acquisition of data for model formation and the PET imaging itself. There is also a fundamental assumption, when using MR data to correct PET, that the reference MR volume is in the same position as the reference PET position. Although there is evidence to suggest that this might be reasonable (end-expiration is the most repeatable position [41]), it ultimately assumes that there are no major deviations (such as those due to bulk motion), whereas direct PET-PET registration methods can avoid this limitation. Evaluation using real PET data would be beneficial at assessing these practical issues and will be investigated in the future.

The IC technique used an indirect correspondence model approach using PET gates as the surrogate. This was compared with a DC motion model approach using an MR-based respiratory signal as the surrogate, but an alternative approach would have been to use a respiratory signal derived from PET data as the surrogate (e.g. [83, 86]). However, such an approach would introduce uncertainty as to how to relate the different (but similar) surrogate signals used to form the model (i.e. MR-based signal) and apply it (PET-based signal). The IC approach has no such problem since the value of the signal used to form the model is optimised based on the richer information contained in the PET gates used as surrogate image volumes.

6.6 Conclusion

The method outlined in this chapter represents a proof-of-principle of a new class of PET motion correction techniques. More complex implementations, using some of the improvements outlined in the previous section, are possible. In addition, this technique is one of the first attempts to use both PET and MR imaging data to estimate motion fields for PET motion correction.

In summary, the technique described represents an important addition to the literature on PET-MR motion correction: it has been shown that good, reliable PET motion correction performance can be achieved without continuous or repeated use of the MR scanner. This potentially makes incorporating motion correction into clinical protocols much more feasible.

The next chapter will show how this constrained registration paradigm can be incorporated into the PET reconstruction itself. This relies upon MCIR-type reconstruction, during which the internal parameter of the motion model will be statistically estimated using the PET data, rather than using images as in the IC approach in this chapter. Note that, since the motion model in the next chapter is also an IC model, the method in this chapter will be referred to as motion-estimating RTA (ME-RTA), whereas the method in the next chapter will be referred to as motion-estimating MCIR (ME-MCIR) for clarity.

Chapter 7

Reduced-Parameter Joint Motion/Image Reconstruction

7.1 Introduction

In the previous chapter, a proof-of-principle was introduced, which used an indirect-correspondence motion model to estimate the motion position of each PET gate in a dataset. One might wish to improve this process by using a single objective function that can be optimised. Since iterative image reconstruction is also an optimisation process, it is tempting to try and unite both motion and activity estimation into the same algorithm. The outline of such a technique will be discussed in this chapter, which will be referred to as motion-estimating MCIR (ME-MCIR).

7.1.1 Motivation

For a given motion estimate, a typical way to motion correct the activity image during reconstruction is with MCIR. As discussed in section 4.5.1, this is achieved by factorising the system matrix into two matrices: one is the original, static projection matrix, $\boldsymbol{A}$, and the other is a motion matrix, $\boldsymbol{M}$. This tells us that in the image reconstruction, the image estimate, $\boldsymbol{f}^{(k)}$, must be first transformed into the position described by the motion estimate prior to projection and comparison with the data.

The ratio sinogram is then back-projected as usual, but then it is transformed back into the motion-free position according to the transpose of $\boldsymbol{M}$. Assuming that the motion estimate is accurate, this procedure should reconstruct an image estimate that is relatively motion-free.

Moving towards an MCIR approach has a number of other benefits. Firstly, MCIR is desirable due to its improved ability to provide quantitatively-accurate results. Moreover, the low SNR of gated sinograms is less of a problem as it will not introduce bias into the reconstruction (a limitation imposed by the non-negativity constraint of iterative reconstruction algorithms [151]). Instead, the full dataset is used each time the image estimate is updated.

Finally, the previous chapter tested motion estimation in RTA, which assumed that the correct attenuation map for each gate was known. This is unrealistic, since having such maps would remove the need to estimate the motion from the PET data in the first place; instead the attenuation maps could be registered. Alternatively, non-attenuation corrected images could be used for estimation, followed by a second reconstruction using attenuation maps transformed by the motion estimates. However, this would be impractical.

The aim of this chapter is to introduce a theoretical framework that derives an expression that will be able to estimate motion, using the same Poisson log-likelihood objective function used to estimate the image. This is to be achieved without using regularisation, to stay as true to the pure maximum likelihood formulation as possible. From a mathematical perspective, $\boldsymbol{M}$ is a sparse matrix constructed according to the motion estimate it represents, $\{\Delta\vec{r}_j\}$. Specifically, each column of $\boldsymbol{M}$ should sum to 1, and the entries which are nonzero are determined by calculating the new (possibly interpolated) position of pixel j.

To keep the following example clear, it will be restricted to nearest-neighbour interpolation, such that each column of $\boldsymbol{M}$ contains only one nonzero entry. This example will be further restricted to a 1-dimensional image. In such a scheme, pixel

j will shift to pixel j', due to a shift Δx_j. This could be identified according to

$$j' = j + \left\lfloor \frac{\Delta x_j}{d_x} \right\rceil \tag{7.1}$$

where d_x is the pixel size (in mm) and $\lfloor \rceil$ denotes rounding to the nearest integer. One might conceivably construct $\boldsymbol{M}$ by setting each element $\left(j, j + \left\lfloor \frac{\Delta x_j}{d_x} \right\rceil\right)$ to 1, and the rest set 0. However, the aim is to optimise an objective function with respect to the motion parameters, $\{\Delta x_j\}$. The most obvious way to achieve this would be through a gradient-based optimisation scheme. In this example, there is no clear way to analytically and directly find a gradient of $\boldsymbol{M}$ with respect to its indices (which, in this formulation, is where the motion parameters feature in the mathematics), and instead one would be restricted to empirical methods of finding a gradient whilst maintaining the matrix formulation of the motion. This motivates the work in this chapter, where the problem is reformulated in a way that the motion parameters are more accessible for direct optimisation.

7.1.2 Previous Joint Motion/Image Estimation Methods

Since the introduction of MCIR, there have been a number of methods incorporating motion estimates into PET image reconstruction. These are usually acquired using other modalities. For example, [126] used tagged MRI, which magnetises a grid into the tissue of the subject. This allows a high quality of motion tracking, especially in homogeneous tissues. However, it does not work well in areas with low signal (such as the lungs), and is generally incompatible with other MR protocols. Tagged MRI was also used in [147, 178].

Methods involving PET-based motion estimates into reconstruction are also possible, although these are usually impractical since they require multiple reconstructions. For example, [115] use FBP to obtain gated SPECT images, from which motion estimates were acquired using optical flow and incorporated into a one-step-late reconstruction via MAP with a temporal Gibbs term.

This impractical 2 reconstruction approach motivates incorporation of the motion estimation procedure into the reconstruction. A number of methods for joint esti-

mation of motion and images have been proposed. In 2003, Jacobson and Fessler set out a theoretical framework for this purpose, based on penalised likelihood [107]. This was formulated with vector of general motion parameters, although this paper did not verify the approach by experiment. Another framework based on a Gaussian mixture model was suggested by [106], although this was only tested in a rigid case.

All tested methods of joint estimation generally involve MAP-type approaches. For example, [116], where the motion estimation was included as a prior in the objective function. However, this method had trouble fixing the associated regularisation parameter. This was addressed later in [179], although long reconstruction times are reported. This method used B-splines, which reduces the number of motion parameters to be estimated (although there will still be many, in general). Other Bayesian approaches include [111], who applied an elastic penalty term to estimate motion in 2D cardiac ECT, and [112], who used their own optimisation 'RM' optimisation algorithm in lieu of EM. This latter method involves a joint objective function with a regularisation term for both the image and motion, and was tested on cardiac SPECT. More recently, a B-spline–based penalised likelihood approach was proposed [97], which takes (possibly misaligned) attenuation into account too.

A known problem with joint estimation algorithms is the increased computational cost [109]. In part, this is due to the significant increase in parameters to be estimated. Some approaches require arbitrary parameterisation of the motion using B-splines to address this, increasing robustness to noise in the motion estimation. Penalisation terms are also required for this purpose. Below, these problems are addressed by incorporating a motion model into a reconstruction algorithm to reduce the parameterisation, allowing use of voxelwise motion fields constrained to physically-realistic positions.

7.2 Background

7.2.1 Reformulating the Data Model

As discussed in the theoretical background in section 2.2.3, iterative image reconstruction works with a model of the mean of the data. This is used to estimate effects which affect the data during acquisition, with the aim of producing a reliable and accurate reconstruction of the activity in the PET FOV. For example, we might factorise the system matrix in (2.10) according to the MCIR formulation of the motion-compensated model,

$$\boldsymbol{q} = \boldsymbol{AMf} \tag{7.2}$$

where $\boldsymbol{q}$ is the mean data sinogram.

However, it is worth noting that there is an implicit step in formulating this discrete data model. The activity distribution is effectively a continuous scalar function in real space, $f(\vec{r})$. As discussed in section 2.1, this continuous distribution is observed by the PET scanner as a series of projections, $p(\vec{u})$, where $\vec{u}$ denotes projection space. Here, $\vec{u}$ will be used to denote functions which are sinogram-type functions explicitly. The geometric transformation from f to p is known as the X-ray transform. The X-ray transform is a Fredholm equation of the first kind, and this means that it has a useful property: if one can express projections, p, as a set of orthogonal eigenfunctions,

$$p(\vec{u}) = \sum_{i=1}^{I} p_i d_i(\vec{u}) \tag{7.3}$$

then we can discretise the transform to recover the matrix equation in (2.10). Generally in image processing, only this discrete representation of an image is considered. However, by appealing back to this implicit discretisation of the transformation, it is possible to retain an explicit link between the underlying continuous distribution of activity, $f(\vec{r})$, and the discrete intensities used to represent it as an image, $\boldsymbol{f}$:

$$f(\vec{r}) = \sum_{j=1}^{J} f_j b_j(\vec{r}) \tag{7.4}$$

where b_j are the image basis functions. These could, in principle, be anything, such as a voxel, box splines, truncated Gaussians, or any other basis function deemed useful for a given problem.

It is feasible, then, to construct a model of the mean of the data, q, which is a continuous version of (7.2):

$$q(\vec{u}, \Delta\vec{\boldsymbol{r}}) = \mathcal{R}\mathcal{T}(\Delta\vec{\boldsymbol{r}})\{f(\vec{r})\} \tag{7.5}$$

where $\Delta\vec{\boldsymbol{r}}$ are some motion parameters, $\mathcal{R}$ is the continuous X-ray transform, and $\mathcal{T}(\Delta\vec{\boldsymbol{r}})$ is an operator which transforms its operand according to the (currently generic) motion parameters, $\Delta\vec{\boldsymbol{r}}$. From here on, $\mathcal{T}$ will be referred to as the 'motion operator'.

7.2.2 Defining the Action of the Motion Operator

The motion operator is currently undefined, and it is worth pausing to consider what action it performs. Suppose that the activity distribution after transformation is denoted $g = g(\vec{r}, \Delta\vec{\boldsymbol{r}})$. For simplicity, we will only consider shift-invariant basis functions $b_j(\vec{r}) \equiv b(\vec{r} - \vec{r}_j)$. This means that each basis function is identical, but supported on a grid defined by basis centres, $\vec{r}_j$. We further assume that $\mathcal{T}$ shifts the centre of each and every basis function by its respective displacement, $\Delta\vec{r}_j$. Therefore,

$$g(\vec{r}, \Delta\vec{\boldsymbol{r}}) = \sum_{j=1}^{J} f_j \mathcal{T}(\Delta\vec{\boldsymbol{r}})\{b(\vec{r} - \vec{r}_j)\} = \sum_{j=1}^{J} f_j b(\vec{r} - \vec{r}_j - \Delta\vec{r}_j). \tag{7.6}$$

There are several points worthy of note here. First, in a strictly formal sense, this set of shifted basis functions, $\{b_j\}$, is no longer guaranteed to span the entire image space. Should one wish to create an image of the motion-transformed image, $\boldsymbol{g}$, one would need to resample according to the original, unshifted basis functions. If the basis functions are not allowed to warp and change shape with the underlying grid, rips and tears can appear in the image. However, since the derivation in this chapter is a means to an end for arriving at a motion estimation algorithm, it is

assumed that application of an image transformation will involve interpolation (this is discussed further in section 7.2.4). A full treatment would involve a set of general basis functions.

The second point to note is that the total intensity in the image might not be preserved by this treatment. A general warped basis function will change in volume, which can be taken into account by normalising it with its new volume. This is known as mass preservation [122]. A practical way to address this was employed in [180], where the volume changes were approximated using tetrahedra.

This rearrange-and-interpolate process is analogous with application of the motion matrix $\boldsymbol{M}$ to a vector of image intensities $\boldsymbol{f}$. Resampling the shifted basis functions $b(\vec{r} - \vec{r}_j - \Delta\vec{r}_j)$ back to a regular grid $b(\vec{r} - \vec{r}_{j'})$ – where j' denotes the shifted basis index – requires calculating the weights generated by any overlap of the old and new basis functions,

$$g_{j'} = \sum_{j=1}^{J} f_j \int_{-\infty}^{\infty} b(\vec{r} - \vec{r}_j - \Delta\vec{r}_j) b(\vec{r} - \vec{r}_{j'}) \, \mathrm{d}^3 r \equiv \sum_{j=1}^{J} M_{jj'} f_j. \tag{7.7}$$

The fact that this expression for calculating the discrete motion matrix drops out of the formulation is reassuring. Note that the calculation of the integral – or approximation of it – is identically the interpolation step performed when transforming a digital image. Note also that this results in a loss of information: interpolation generally results in superposition of the original image intensities due to the overlaps. The interpolated image will not (unless by design) contain information on the components of this superposition, and is therefore irreversible. Due to this, $\boldsymbol{M}$ is not necessarily an orthogonal matrix, possibly creating problems in any algorithm where a motion matrix formed from inverted motion fields is substituted for $\boldsymbol{M}^{\mathrm{T}}$.

Secondly, this point is where many previous methods of joint activity/motion estimation stop: one tries to empirically estimate $\Delta\vec{r}_j$ through some means. Of course, to do so with an image volume generates a very large number of additional degrees of freedom to an already ill-posed estimation problem. At one extreme is voxelwise motion estimation, in which one must estimate J image intensities, and a further $3JG$ motion parameters (i.e. the (x, y, z) displacement of each voxel, for each of

a total of G gates). It is clear that this would take a long time to calculate, and could encounter issues due to the poor qualities associated with PET data, such as low SNR. Instead, it is common to reduce the dimensionality of the problem by, for example, only looking at displacements of control points of a B-spline grid (e.g. [179]). Other methods might seek to increase the amount of information available to the solver by including data from MR, or perhaps try to minimise impact of noise through regularisation [97].

Instead, what is proposed here is the recognition that the displacements $\Delta\vec{\boldsymbol{r}}$ can be estimated with a parameterised motion model: $\vec{\boldsymbol{\phi}}(\boldsymbol{s})$. Therefore,

$$g(\vec{r}, \boldsymbol{s}) = \sum_{j=1}^{J} f_j b(\vec{r} - \vec{r}_j - \vec{\phi}_j(\boldsymbol{s})) \tag{7.8}$$

is the result of applying the motion operator to continuous activity distribution. The operator, in this case, is now a function of $\boldsymbol{s}$, rather than the full $\Delta\vec{\boldsymbol{r}}$. This description of g is a general one since there are no assumptions about the type of motion model used, which parameters it takes as argument, nor which basis is used to describe the image. In the following, we will choose to use the motion model used in the previous chapter: one which only depends on a scalar internal variable, s. This reduces the dimensionality of the joint estimation of activity and full voxelwise motion estimation from $J + 3JG$ to $J + G$.

7.2.3 Finding a Motion Estimation Equation

As mentioned above, PET data exhibit noise of a Poisson nature. For a set of I mutually-independent Poisson variables, the probability of observing all measurements $\boldsymbol{m}$ with respective mean outcomes $\boldsymbol{q}$ is

$$P(\boldsymbol{m}|\boldsymbol{q}) = \prod_{i}^{I} \frac{q_i^{m_i} \exp(-q_i)}{m_i!} \tag{7.9}$$

This equation can be used to statistically model the mean of the variables $\boldsymbol{q}$ *given* measurements $\boldsymbol{m}$, such that, with these measurements, the likelihood of model $\boldsymbol{q}$ is $L(\boldsymbol{q}|\boldsymbol{m})$. Further to this, we define L such that it is the natural logarithm of

P. This is a permissible change since the logarithm is a monotonically-increasing function that will exhibit maxima in the same locations as the original function. The advantage is that it converts the product into a sum. Therefore,

$$L(\boldsymbol{q}|\boldsymbol{m}) = \sum_i (m_i \log(q_i) - q_i - \log(m_i!)) \tag{7.10}$$

which is just the Poisson log-likelihood function referred to in section 2.2.3.

The overall aim of this approach is to estimate the scalar motion parameter s, which drives the distribution of weights in $\boldsymbol{M}$, or, more usefully, the transformation imposed by operator $\mathcal{T}(s)$. As discussed above, a method to optimise the indices of a matrix with respect to a continuous parameter is not apparent. Instead, we will work with the continuous model of the data, which will eventually need to be discretised with a sampling equivalent to $\boldsymbol{m}$. To make this apparent in equations, a square bracket with a subscript will denote a 'discretisation bracket', to show that this operation is yet to be executed. Therefore, note that in this case $q_i \equiv [q(\vec{u}, s)]_i$. In other words, q_i constitute a discrete set of values which are samples of the continuous function $q(\vec{u}, s)$, discretising it. Note that this is simply a generalisation of the notation $[\boldsymbol{v}]_i$ for the i^{th} element of a vector, which in general is referred to by $\boldsymbol{v}$.

Therefore, we can rewrite (7.10)

$$L(\boldsymbol{q}|\boldsymbol{m}) = \sum_i \left(m_i \log_{\mathrm{e}} \left([q(\vec{u}, s)]_i\right) - [q(\vec{u}, s)]_i - \log(m_i!)\right) \tag{7.11}$$

There are many ways to optimise this equation with respect to s, but for simplicity the gradient ascent method will be used. Given that a gradient of L with respect to s exists, it is possible to iteratively estimate a maximal solution of s, $\hat{s}$, according to the update equation

$$s^{(k+1)} = s^{(k)} + \beta^{(k)} \frac{\mathrm{d}L(s|\boldsymbol{m}, \boldsymbol{f}^{(k+1)})}{\mathrm{d}s}\bigg|_{s=s^{(k)}}. \tag{7.12}$$

Note that this requires the updated estimate of the image, $\boldsymbol{f}^{(k+1)}$, which is assumed static during the motion parameter update. This means that this is an alternating scheme, where the image and the motion parameters are updated in turn during

each iteration of the main reconstruction.

As (7.12) suggests, a derivative of L with respect to s is required. Due to the work in sections 7.2.1 and 7.2.2, there is an explicit relationship between q and s. First, recall that for any function $f(x)$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\log_{\mathrm{e}}(f(x)) = \frac{f'(x)}{f(x)}, \tag{7.13}$$

so that we can take the derivative of (7.11) to get

$$\frac{\mathrm{d}L(s|\boldsymbol{m}, \boldsymbol{f}^{(k)})}{\mathrm{d}s} = \sum_{i=1}^{I}\left(\frac{m_i}{[q(\vec{u},s)]_i} - 1\right)\left[\frac{\mathrm{d}q(\vec{u},s)}{\mathrm{d}s}\right]_i. \tag{7.14}$$

where, short of discretisation, the first term in parentheses in the sum is already known. However, the derivative of the model of the mean with respect to s still needs to be calculated.

First, consider (7.5) and its dependence upon s. This has been modelled such that the projection and motion operators are separate. Due to this, $\mathcal{R}$ is explicitly not a function of s, since all s-dependence is carried by $\mathcal{T}$. Therefore,

$$\frac{\mathrm{d}q(\vec{u},s)}{\mathrm{d}s} \equiv \frac{\mathrm{d}\mathcal{R}\mathcal{T}(s)\{f(\vec{r})\}}{\mathrm{d}s} = \mathcal{R}\frac{\mathrm{d}g(\vec{r},s)}{\mathrm{d}s}. \tag{7.15}$$

As already described in (7.8), $g(\vec{r},s)$ is simply the image after application of the motion transformation: all J of the basis functions are displaced according to the motion transformation described by the motion model $\vec{\boldsymbol{\phi}}(s)$. Explicit s dependence in g is carried by the shifted basis functions, $b(\vec{r} - \vec{r}_j - \vec{\phi}_j(s))$. In 1D, using the substitution $v_j = x - x_j - \phi_{jx}(s)$ allows us to find the derivative of this function using the chain rule:

$$\frac{\mathrm{d}b(x - x_j - \phi_{x,j}(s))}{\mathrm{d}s} = \frac{\mathrm{d}b(v_j)}{\mathrm{d}v_j}\frac{\mathrm{d}v_j}{\mathrm{d}\phi_{jx}}\frac{\mathrm{d}\phi_{jx}(s)}{\mathrm{d}s} \equiv -\frac{\mathrm{d}b(v_j)}{\mathrm{d}v_j}\frac{\mathrm{d}\phi_{jx}(s)}{\mathrm{d}s}. \tag{7.16}$$

The extension beyond 1D requires the equivalent of the chain rule for scalar functions which take vector arguments. We now define $v_{j,k} = r_k - r_{jk} - \phi_{jk}(s)$ for each cardinal

direction, $r_k \in \{r_1, r_2, r_3\}$, the derivative in the k^{th} direction is

$$\left[\frac{\mathrm{d}b(\vec{r}-\vec{r}_j-\vec{\phi}_j(s))}{\mathrm{d}s}\right]_k = \frac{\partial b(v_{j1}, v_{j2}, v_{j3})}{\partial v_{jk}} \frac{\partial v_{jk}}{\partial \phi_{jk}} \frac{\mathrm{d}\phi_{jk}(s)}{\mathrm{d}s}. \tag{7.17}$$

The full 3D derivative is then just the sum over each direction, k,

$$\frac{\mathrm{d}b(\vec{r}-\vec{r}_j-\vec{\phi}_j(s))}{\mathrm{d}s} = -\sum_{k=1}^{3} \frac{\partial b(v_{j1}, v_{j2}, v_{j3})}{\partial v_{jk}} \frac{\mathrm{d}\phi_{jk}(s)}{\mathrm{d}s}. \tag{7.18}$$

A convenient and intuitive shorthand for this can use the gradient operator, $\vec{\nabla}$, defined as

$$\vec{\nabla} f(r_1, r_2, r_3) = \vec{e}_1 \frac{\partial f}{\partial r_1} + \vec{e}_2 \frac{\partial f}{\partial r_2} + \vec{e}_3 \frac{\partial f}{\partial r_3} \tag{7.19}$$

where e_k are the normalised Cartesian basis vectors. Notice that $\vec{\nabla}$ is a vector, as is $\vec{\phi}$ (and its subsequent derivative with respect to s). Therefore (7.18) can be considered a dot product between these two terms:

$$\frac{\mathrm{d}b(\vec{r}-\vec{r}_j-\vec{\phi}_j(s))}{\mathrm{d}s} = -\vec{\phi}'_j(s) \cdot \vec{\nabla}_{\vec{v}_j} b(\vec{v}_j), \tag{7.20}$$

where the prime denotes total derivative and the subscript $\vec{v}_j$ on the gradient operator makes the derivative variables explicit. Combined with the definition for $g(\vec{r}, s)$, the derivative is now well defined:

$$\frac{\mathrm{d}g(\vec{r}, s)}{\mathrm{d}s} = -\sum_{j}^{J} \vec{\phi}'_j(s) \cdot \vec{\nabla}_{\vec{v}_j} b(\vec{v}_j)\, f_j. \tag{7.21}$$

Note that, so far, this result is completely general; there are no assumptions on the image basis functions used, nor the dependence of $\vec{\phi}$ upon s.

The result in (7.21) should be somewhat intuitive since it shows that two parameters are emphasised when calculating the derivative of the log-likelihood with respect to s. First, there is the derivative of the motion model, $\vec{\phi}'(s)$. Since $\vec{\phi}$ can be considered a displacement that depends on some phase-like[1] parameter s, its derivative with

[1]Note that, since ϕ is currently a general motion model, s is *not* actually phase-like unless formulated as such. For example, in the work of this thesis, s is an amplitude: the head-foot displacement of the apex of the right hemidiaphragm. The phase-like quality referred to here is primarily an analogy, corresponding to the cyclic nature of the displacements due to respiratory motion.

respect to this parameter, ϕ', is something akin to a velocity. This will lend increased sensitivity to regions of the image which move 'faster'. Secondly, there is a space-like derivative of the basis functions. This, coupled with the intensities f_j, emphasises any inhomogeneous features in the image, such as high-contrast edges.

Gathering all these pieces together, we now have a complete expression for the derivative of the log-likelihood. Equation (7.14) can now be written

$$\frac{\mathrm{d}L(s|\boldsymbol{m}, \boldsymbol{f}^{(k)})}{\mathrm{d}s} = \sum_{i=1}^{I} \left(\frac{m_i}{[\mathcal{R}\{g(\vec{r}, s)\}]_i} - 1 \right) \left[-\mathcal{R} \left\{ \sum_{j=1}^{J} \vec{\phi}'_j(s) \cdot \vec{\nabla}_{\vec{v}_j} b(\vec{v}_j) f_j \right\} \right]_i . \quad (7.22)$$

The meaning of this gradient is now apparent, with the weights (velocity, intensity, and edges) that prioritise the gradient on the far right, and a 'switch' before that, which turns off the update gradient when the image converges.

Note that in this chapter, the data is indirectly driving the model since it is a set of variables estimated using PET sinogram data. Therefore, this can also be considered a type of indirect correspondence (IC) motion model, and is the first example of one being driven using PET sinograms as the motion surrogate.

7.2.4 Discretising the Update Equation

A problem still exists however: the model of the mean is still in a continuous formulation, and must be discretised. An amount of approximation will be necessary here, as would be the case for any conversion from continuous space to discrete space.

Firstly, and most straightforward, is the discretisation of the model of the mean, $q(\vec{u}, s)$. This is, by definition, just the discrete equivalent of the model, or,

$$q_i \equiv [q(\vec{u}, s)]_i = [\boldsymbol{AMBf}]_i \quad (7.23)$$

where, as before f_j are the image intensities. In addition, there are 3 matrices: the system matrix $\boldsymbol{A}$, defined in (2.9), $\boldsymbol{M}$, which represents the rearrangement of the basis functions and resampling implicit in $\mathcal{T}(s)$, and $\boldsymbol{B}$, a matrix which carries the information about the basis functions, their sampling, and how they overlap. In the

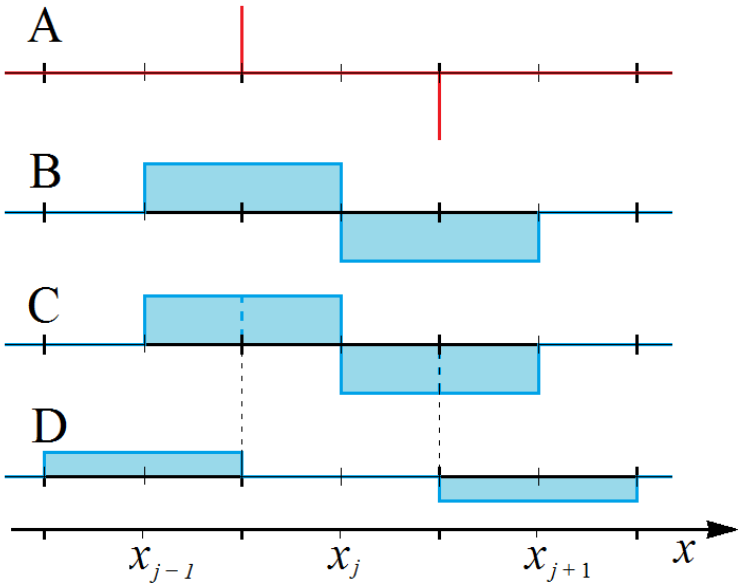

Figure 7.1: Given the approximation in (7.25), we can represent the Dirac distributions, (A), as unit-height top-hat functions centred at $x_j \pm {}^1\!/_2$, (B). To resample these back onto the original grid, we notice that the halves in a pixel centred at x_j cancel, leaving only half in each of x_{j-1} and x_{j+1} respectively (C). The result is two top-hat functions adjacent to the central pixel, with half-unit height.

case of pixels or voxels, $\boldsymbol{B} \to \boldsymbol{I}$, where $\boldsymbol{I}$ is the identity matrix, recovering (7.2).

Next, the rightmost term in (7.22) must be expanded. For concreteness, voxels will be used as the image basis functions, although this method should still work for others. The $\vec{\nabla}$ operator acts upon each spatial component of b separately. Since the voxel basis function $\Pi(\vec{r})$ is separable, we only need to consider the derivative of the 1D top-hat function, which comprises the 3D analogue. The derivative of the top-hat function is

$$\frac{\mathrm{d}\Pi(x)}{\mathrm{d}x} \equiv \frac{\mathrm{d}}{\mathrm{d}x}\left\{H\left(x+{}^1\!/_2\right) - H\left(x-{}^1\!/_2\right)\right\} = \delta\left(x+{}^1\!/_2\right) - \delta\left(x-{}^1\!/_2\right) \tag{7.24}$$

where H and δ denote the Heaviside step- and Dirac delta functions respectively. According to [181], a possible discrete representation of the Dirac delta function is

$$\delta(x) = \lim_{\epsilon\to 0} \frac{1}{\epsilon}\Pi\left(\frac{x}{\epsilon}\right) \tag{7.25}$$

where, since we are confined to a uniform grid, the $\pm{}^1\!/_2$ in (7.24) leads to top-hat functions displaced from the grid by half the grid distance. On resampling and

interpolating (refer to figure 7.1 for details), we find

$$\frac{\mathrm{d}\Pi(x)}{\mathrm{d}x} \approx \frac{1}{2}\big(\Pi(x+1) - \Pi(x-1)\big), \tag{7.26}$$

which is the central difference of the top-hat function. The central difference in a given direction can be expressed as a matrix, $\boldsymbol{D}$, which will generally be constructed from off-diagonal ± 1s, all divided by 2. The generalisation into 3D should be apparent, with the exact ordering of these elements defined by the number of voxels in each direction of the image. Therefore, for a 3D derivative of an image, we have a set of 3 central difference matrices, $\{\boldsymbol{D}_x, \boldsymbol{D}_y, \boldsymbol{D}_z\}$, which denoted $\vec{\boldsymbol{D}}$ in shorthand.

Above, the gradient operator and the derivative of the motion model constitute a scalar product. Discretising $\vec{\phi}'$ should be straightforward, since it will simply be sampled at the grid points, to obtain $\vec{\boldsymbol{\phi}}'$. The scalar product can be expressed in a linear fashion by supporting each component of $\vec{\boldsymbol{\phi}}'$ on a diagonal matrix, such as $\mathrm{diag}\,(\boldsymbol{\phi}'_x)$. As such, the values of the derivative motion model behave as weights for each direction of $\vec{\boldsymbol{D}}$, so that we can define the overall matrix

$$\boldsymbol{D}_{\mathrm{W}} = \sum_{k=1}^{3} \mathrm{diag}\,(\boldsymbol{\phi}'_k)\,\boldsymbol{D}_k, \tag{7.27}$$

which will be referred to as the weighted derivative matrix. Notice that $\boldsymbol{D}_{\mathrm{W}}$ is once again a regular $[J \times J]$ data matrix; in a spatial sense, it is a scalar.

Finally, we can express the discretised derivative of q with respect to s as

$$\left[\frac{\mathrm{d}q}{\mathrm{d}s}\right]_i = [\boldsymbol{A}\boldsymbol{D}_{\mathrm{W}}\boldsymbol{M}\boldsymbol{f}]_i \tag{7.28}$$

where $\boldsymbol{D}_{\mathrm{W}}$ is defined in (7.27). This then allows us to fully discretise (7.22):

$$\frac{\mathrm{d}L(s|\boldsymbol{m}, \boldsymbol{f}^{(k)})}{\mathrm{d}s} = -\left(\boldsymbol{A}\boldsymbol{D}_{\mathrm{W}}\boldsymbol{M}\boldsymbol{f}\right)^{\mathrm{T}} \left(\frac{\boldsymbol{m}}{\boldsymbol{A}\boldsymbol{M}\boldsymbol{f}} - \mathbf{1}\right). \tag{7.29}$$

Notice that the quantity on the left hand side is a scalar value, as should be expected. As a gradient, this equation can be used to iteratively update estimates of the respiratory signal for given PET data, which is also a scalar.

7.3 Methods and Materials

The aim for the rest of this chapter is to assess the efficacy of joint motion and activity estimation during reconstruction using the update formula derived above. First the reconstruction will be tested using emission-only simulations of 4 volunteers with lesions in 6 positions each. The resulting images will be tested against motionless simulations and compared with uncorrected images and the ME-RTA method introduced in the previous chapter. Second, the effect of increasing the number of gates will be tested. This will be compared against a PET to PET direct registration RTA technique to see how both methods cope with the increased noise levels in the gated data. This requires more careful simulation of the intra-gate motion, which will be explained in more detail in section 7.3.2.

7.3.1 MR Data

Although the method of simulation in this chapter is slightly different, the same MR data was used to generate the PET data. This involved a dynamic (i.e. multiple 3D in sequence) scan for generating motion estimates and a UTE scan. The motion estimates were split into two sets, one for including motion into the simulations, and the other for building the parameterised motion model. The UTE images were used to generate emission and attenuation maps for the simulation process. Acquisition details for the MR data is described in section 6.4.1. These were registered using lreg, as before, using the parameters specified in section 6.4.3 [173].

7.3.2 Data Simulation

The PET data used for the experiments in this chapter are similar in design to those simulated in chapter 6, but differ in implementation: MATLAB was used to simulate the PET, rather than STIR as before. This was predominantly for convenient flexibility in development of the reconstruction algorithm and was facilitated by the provision of MATLAB-based projectors provided by Andrew Reader. A detailed discussion of this is presented in chapter 5.

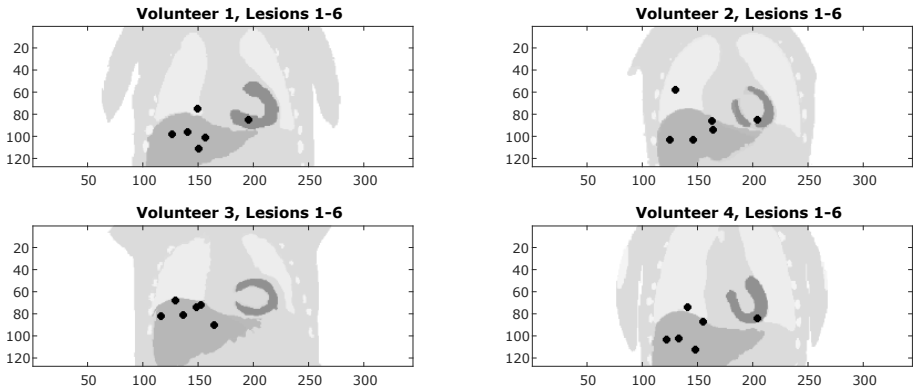

Figure 7.2: **Relative lesion locations projected onto coronal slice** This figure shows where each of the 6 lesions were positioned for each volunteer, projected onto a coronal slice. These positions were selected due to the higher than average displacement observed in the motion fields. Note that the tissue type of the lesion in these images might not reflect the true tissue type in which it is placed.

The PET data were designed using the same pipeline: the UTE data were segmented into tissues which are relevant to both [^{18}F]-FDG tracer uptake and attenuation, generating an image for each. Lesions were artificially segmented into the FDG images, although the method of doing this varied slightly to the previous chapter and will be detailed below. As before, motion was simulated by transforming the FDG images into a range of respiratory positions measured using dynamic volunteer MRI data. This provided the motion of each PET gate. The motion was calculated using non-rigid registration software (see section 6.4.3 for details). Finally, the transformed FDG images were projected to create noiseless FDG sinograms. Poisson noise was simulated to the level of SNR expected for 50 million counts, divided across the gates according to the period each volunteer would have spent in each gate according to their observed breathing pattern. Only a single realisation of noise was simulated for each gate. It is these noisy sinograms that were reconstructed. Note that attenuation was *not* included in these experiments: this will be addressed in the next chapter.

Differences in Simulation Design

As mentioned, there were minor differences in the generation of the PET simulations used in this chapter. These differences were introduced to try to improve on the experimental design used in the previous experiments.

Firstly, a skew was observed in the results of the previous chapter. This was thought to be due to poor lesion placement: lesions placed in regions of less displacement would exhibit little degradation and therefore relatively little could be done to improve them. This introduced a systematic error into the resulting averages, resulting in an apparently good performance of using no motion correction. Similarly, the ME-RTA method appeared to recover lesion uptake on average, but performance was poorer in the lungs.

To address this issue, more thought was put into lesion placement. For each volunteer, a maximum displacement motion field was generated, comprised from the magnitude of the motion field corresponding to the largest intake of breath (recall from before that the reference position was the end-expiration). In this magnitude motion-field–like image volume, high intensity corresponded to areas of high displacement. Therefore, motion 'hotspots' were identified and lesions were placed in these positions. A disadvantage of this approach is that there is no regularity in lesion positioning, making it harder to present the placements of all lesions here. The relative positioning of the lesions, as projected onto a coronal slice as a guide, are presented in figure 7.2. Care was taken to place lesions in a wide variety of tissues and regions of the body.

A total of 6 lesions were simulated per volunteer. These were placed in regions of high motion to avoid the issues with relatively-static lesions observed in the previous chapter. These lesions were simulated as spherical, with a diameter of 12 mm. This time contrast, rather than peak uptake, was kept constant, such that each lesion had a 3:1 contrast with the background tissue (note that this is not reflected in the lesions displayed in figure 7.2). The SUVs used for different tissues are tabulated in table 5.1.

Another difference in the simulation was an improved attempt to model intra-gate motion, since this would be a leading source of error when testing gate-based motion correction in a real scenario. Additionally, an experiment was planned for this chapter (see section 7.3.6, below) which tested the behaviour of ME-MCIR with varying gate number.

To simulate intra-gate motion, each motion field was binned according to which gate

it would contribute towards. This was done by assigning the navigator measurements into G bins, where G is the number of gates to be simulated. The transformed FDG images were then generated as above, but for each and every motion field in that gate-bin (for a visual representation of this, see how the histogram of amplitudes is distributed in figure 4.1: the histogram is not uniform). This set of transformed FDG images was then averaged, creating the final FDG image for projection. It is worth noting that this method is not perfect; intra-gate motion would show a continuous range of motions, and only relatively few dynamic MR images were available (18 for each volunteer), so this approach, whilst an improvement, represents only an approximation to the true continuous range of deformations.

7.3.3 Motion Model Formation

The same motion model was used as before: a 3D non-rigid 2^{nd}-order polynomial function of the navigator signal was formed for each voxel, in each direction. The motion fields used to form the motion model were a separate set to those used to simulate motion in the PET gates, although they were acquired in the same dynamic scan. The details of motion model formation are identical to those in section 6. The motion fields that contribute to each motion model (for a given number of gates) varied slightly due to the way in which motion fields were assigned between simulation and motion model formation. Precedence was given to assigning motion fields to PET simulation, so in some cases the motion model might not have had an equivalent motion field sampled for the range of motion in a given gate.

7.3.4 Defining a Reconstruction Scheme

In order to apply the gradient defined in section 7.2.4, it must be incorporated into a reconstruction scheme. Regular MCIR (with pre-defined motion estimates) can be applied to G respiratory gates to obtain a new image estimate $\boldsymbol{f}^{(k+1)}$ from the

k^{th} image update, $\boldsymbol{f}^{(k)}$, using the iterative update equation

$$\boldsymbol{f}^{(k+1)} = \frac{\boldsymbol{f}^{(k)}}{\sum_{g=1}^{G} \boldsymbol{M}_g^{\mathrm{T}} \boldsymbol{A}^{\mathrm{T}} \mathbf{1}} \sum_{g=1}^{G} \boldsymbol{M}_g^{\mathrm{T}} \boldsymbol{A}^{\mathrm{T}} \frac{\boldsymbol{m}_g}{\boldsymbol{A}\boldsymbol{M}_g \boldsymbol{f}^{(k)}} \ . \tag{7.30}$$

Note that $\boldsymbol{M}_g$ is shorthand for the motion matrix, $\boldsymbol{M}$, formed using the motion estimates resulting from the current output of the motion model, $\vec{\boldsymbol{\phi}}\left(s_g^{(k)}\right)$, where $s_g^{(k)}$ is the k^{th} estimate of the motion parameter s for gate g.

To combine the image and motion parameter estimation during reconstruction an alternating approach was used, which interleaves successive steps of known-motion MCIR (as above) and a gradient-ascent update of the respiratory signal estimate for each gate:

$$s_g^{(k+1)} = s_g^{(k)} + \beta^{(k)} \left(\frac{\boldsymbol{m}_g}{\boldsymbol{A}\boldsymbol{M}_g \boldsymbol{f}^{(k+1)}} - \mathbf{1} \right)^{\mathrm{T}} \boldsymbol{A} \left(\vec{\boldsymbol{\phi}}' \left(s_g^{(k)} \right) \cdot \vec{\boldsymbol{D}} \right) \boldsymbol{M}_g \boldsymbol{f}^{(k+1)}. \tag{7.31}$$

This gradient-ascent scheme uses an empirical stepsize parameter, β, which can be made to vary with iteration number, k. In effect, this scheme can be seen as taking perpendicular, ascending steps in the combined dataspace of the log-likelihood of the image and the respiratory signal estimates with respect to the acquired data.

As described above, the motion model in this work was chosen to be a voxelwise, 2^{nd}-order polynomial function of the motion parameter, s. This means that, for the j^{th} of J voxels,

$$\vec{\phi}_j(s) = \vec{a}_j s^2 + \vec{b}_j s + \vec{c}_j, \tag{7.32}$$

where $\{\vec{a}_j, \vec{b}_j, \vec{c}_j\}$ is a set of polynomial coefficients, containing 9 scalar values in total (3 for the polynomial, in each of 3 spatial directions). It is these coefficients which carry the constraint information set by the MR-derived motion fields. As such, the derivative of (7.32) is

$$\vec{\phi}'_j(s) = 2\vec{a}_j s + \vec{b}_j, \tag{7.33}$$

which, given an input estimate of s, will output the coefficients for the derivative matrix in (7.31) to form the weighted derivative matrix, $\boldsymbol{D}_{\mathrm{W}}$, seen in (7.27).

An example of how each of the variables are combined in a reconstruction pipeline is illustrated in figure 7.3. The initial estimate each image intensity was set to ones, whereas the initial estimate for each of the internal parameters was set to the same value: the midpoint of the range of observed navigator values used to build the model. A total of 40 iterations were used, although the motion update was not started until the 6^{th} iteration. This is to stabilise the initial trajectory of the motion gradient: early image updates mainly consist of low spatial frequencies and are blurred. Since no real motion information can be observed in these early iterations, it was observed that the quality of the motion estimation could be affected.

The stepsize was set such that the mean of first set of absolute gradient estimates was normalised to one millimetre per iteration. This was to ensure that the gradient was at a reasonable scale relative to the first update. No other adjustments were made to the stepsize once it was set. Finally, a VOI was placed over the right lung and liver, similar to the ME-RTA experiments, since it was found to improve estimation accuracy. However, a consequence of this is that reconstruction optimises two different objective functions, since only a subset of projections are used to estimate the motion.

It is worth mentioning that there is significant potential for algorithm optimisation here, since each motion-transformed image must be forward projected both for image estimation and motion estimation. Due to the nature of this method, the interleaving of the updates also means that one must take care to use the correct variables between each substep.

7.3.5 Comparison to ME-RTA

The ME-MCIR technique was tested and compared with the ME-RTA technique, with the uncorrected MLEM reconstruction being used as an example of reconstruction without motion correction. Motion-free data was used to set a gold-standard against which the success of each of these methods was quantified.

Motion-estimating MCIR was used for reconstructing simulated data binned into 6 gates. A separate experiment assessing the behaviour of the algorithm over varying

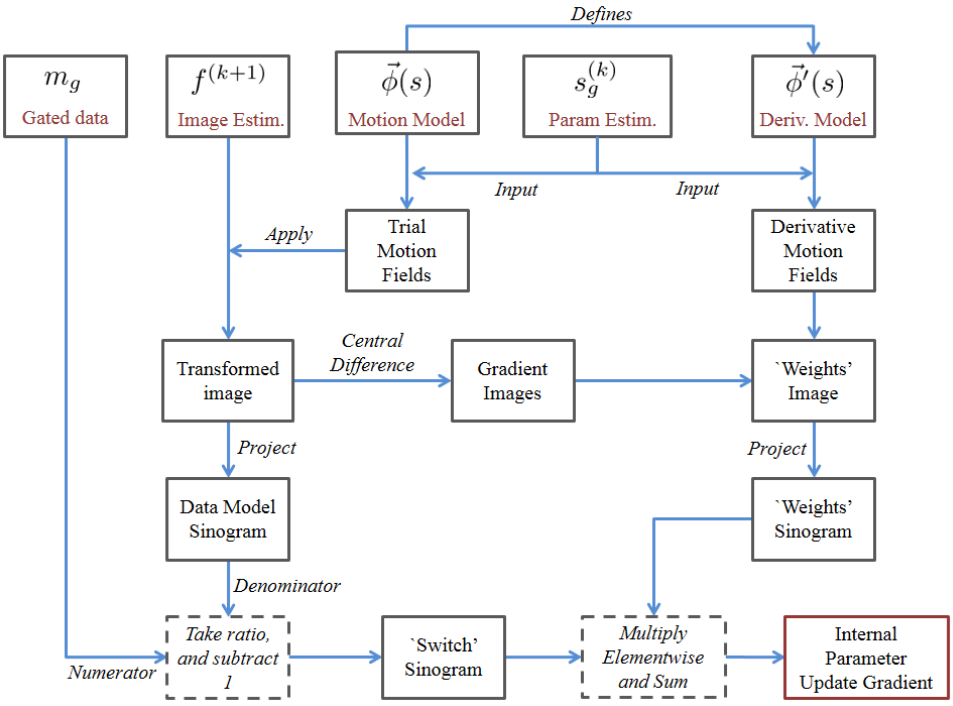

Figure 7.3: This flowchart demonstrates the workflow for finding the gradient of the log-likelihood with respect to the hidden variable. The top row is comprised of the data required to produce the gradient, with successive boxes specifying derived parameters. These continue to combine and interact via the instructions labelled over the arrows. Boxes with a dashed border signify more specific processes. The final result is a gradient for improving the current estimate of the internal parameter. This is applied using gradient ascent to update $s_g^{(k)}$.

numbers of gates is described in section 7.3.6. This was tested over 6 lesion positions in the liver, using data from the same 4 volunteers used in the ME-RTA chapter above. This means that there were 24 datasets, each containing 6 gates, leading to 144 3D data sinograms. There was 1 reconstruction for each of ME-MCIR, uncorrected MLEM, and motionless MLEM, and 6 reconstructions (1 per gate) for ME-RTA, leading to a total of 9 reconstructions per dataset. Therefore this experiment involved a total of 216 reconstructions. In terms of processing speed, preliminary experiments suggest that ME-MCIR will take longer than ME-RTA due to the number of additional projections required for the motion estimation procedure. This will be compared below.

In the previous chapter, all techniques involved reconstruction of individual PET gates (including the motion-free data) in an attempt to avoid any systematic differences introduced by variation in SNR. In this experiment, however, only the ME-RTA method is reconstructed in this fashion. Comparatively, the motionless and uncorrected datasets were each combined into single sinograms prior to reconstruction. The ME-MCIR algorithm being tested was reconstructed using individual gates, but only resulted in one image.

All reconstructions were performed using my own code written in MATLAB using the Siemens mMR projectors described and discussed in chapter 5. Estimation of the respiratory signals was not started in the ME-MCIR reconstruction until after 6 iterations of image reconstruction, since preliminary experiments suggested that starting too early can affect the estimation accuracy. This is thought to be due to early iterations producing images with low spatial resolution, and thus any useful motion information in the image is obscured, drawing the optimisation path into a local maximum.

7.3.6 Varying Gate Number Experiments

A possible advantage of the ME-MCIR technique is that it only needs to estimate 1 additional parameter per gate. This means that there are $J + G$ parameters to be estimated in total, where J is the number of intensity values and G is the number

of gates. This contrasts with other joint-estimation methods, which usually contain many more degrees of freedom. In the extreme case of unconstrained voxelwise estimation, $3JG$ transformation parameters (3 directions for each pixel in each gate) would need to be estimated in addition to J image intensities from an already ill-posed problem. This can be mitigated by introducing prior information, such as in a MAP-based method, or by constraining the motion estimation in some way, often using spline interpolation. This is discussed in more detail in section 4.3.3.

This drastic reduction in parameters might prove to give this joint reconstruction method some robustness to noise. This was tested by varying the number of gates in the reconstruction whilst keeping the total number of counts constant. If it can be shown to be reliable for relatively high numbers of gates (i.e. lower SNR gates), this might prove advantageous for improved motion correction due to a reduction in intra-gate motion.

Totals of 3, 6, 9, and 12 gates were tested. However, due to the long time this experiment takes to run, only 5 datasets were reconstructed (1 dataset from volunteers 1, 2, and 3, and 2 datasets from volunteer 4). The motion-free sinograms were reconstructed with MLEM to set a gold standard. Each gate was also reconstructed separately using MLEM. These separate gate images were then used to directly estimate motion using [173] to provide comparison to a purely PET-based approach via RTA. This formed the benchmark for assessing registration behaviour as gate number (and therefore noise level in each gate) increased.

7.4 Results

As described above, two sets of experiments were performed for this chapter: assessing the performance of the proposed method against other techniques, and observing the behaviour as the number of gates increases. The results for these experiments will be explained and discussed below.

This time, care was taken to make the noiseless equivalents available for use in analysis and lesion detection. The estimated respiratory signal values were used to

create motion transformations for the original FDG maps used in the simulation. This allowed noiseless versions of each reconstructed image to be created (at least from a motion point of view; there are additional effects due to noise and convergence), which could then be used for basic background subtraction, improving lesion delineation. Three-dimensional volumes of interest (VOIs) were placed over each background-subtracted lesion. This aided detection of fainter lesions in the image during the automated data analysis. Segmentation was performed manually using ITK-SNAP [167].

A Comment on Statistical Analysis in this Section

Some tests were performed to ensure the correct statistics were used to assess the results in this experiment. Namely, whether the recovery values were from a normal distribution was assessed using both skew calculations and the Shapiro-Wilk test.

Skew was defined for a set of measurements, $\boldsymbol{x}$, such that

$$S(\boldsymbol{x}) = \frac{3\left(\text{mean}(\boldsymbol{x}) - \text{median}(\boldsymbol{x})\right)}{\sigma(\boldsymbol{x})}, \tag{7.34}$$

where σ is the standard deviation of $\boldsymbol{x}$ [182]. If the absolute skew, $|S|$, of all three methods of dealing with motion (ME-MCIR, ME-RTA, and uncorrected) for a given measurement type (such as peak uptake recovery) were calculated to be less than or equal to 1, data were assumed to be normal.

The null hypothesis in the Shapiro-Wilk test is that the measurements are sampled from a normal distribution. Therefore, a normal dataset will be unable to reject the null hypothesis, according to the test statistic and its associated p-value. P-values below 0.05 were counted as a rejection of the null hypothesis. This was calculated using MATLAB community–sourced code written by user Ahmed BenSaïda.

The result of using these tests on each set of comparative measurements is shown in table 7.1. If no skew was observed and the Shapiro-Wilk test's null hypothesis could not be rejected then mean and standard deviation were used, and significance was tested using a two-tailed paired Student's t-test. Otherwise the median was used

Measurement Dataset	Skewed?	Normal? (p-value)
Lesion Uptake Recovery		
ME-MCIR	No	Yes ($p = 0.076$)
ME-RTA	No	Yes ($p = 0.895$)
Uncorrected	No	Yes ($p = 0.647$)
Lesion Width Recovery		
ME-MCIR	Yes	No ($p = 0.0320$)
ME-RTA	No	No ($p = 0.0218$)
Uncorrected	Yes	No ($p = 0.0311$)
Lesion Position Recovery		
ME-MCIR	No	No ($p < 0.001$)
ME-RTA	No	Yes ($p = 0.098$)
Uncorrected	No	No ($p = 0.039$)

Table 7.1: **Using statistical tests to identify measures of central tendency** This table summarises the results for the tests used to identify the best methods of central tendency for each dataset. Whether the data showed significant skew is shown in the central column, and whether the Shapiro-Wilk test implied a normal distribution is shown in the right hand column. If no skew was measured and the Shapiro-Wilk p-value was greater than 0.05, the data were assumed normal. All tests are based on measurements of 20/24 lesions, where 4 outliers were omitted. This is explained in the text.

instead, alongside interquartile range. In this case, a two-tailed paired Wilcoxon signed rank test was used to test significance in these skewed cases. For both the t-test and the signed rank test, $p = 0.01$ is taken to be a significant result.

Out of 24 lesions assessed, 4 were assigned as outliers: lesion 5 from volunteer 1, lesions 4 and 5 from volunteer 3, and lesion 5 from volunteer 4. This is due to the lesions being undetectable in the uncorrected images, even with the help of manual delineation. Generally the correction methods make the lesions visible again, but detectability was not formally assessed. All 4 of these lesions were lung lesions, located slightly above the diaphragm. Poor detection was due to the contrast of lung lesions not being high enough to use the measurement techniques effectively over the background noise. Whilst these could have still been used to calculate averages for the correction techniques, it would have required use of unpaired statistical tests and was deemed undesirable.

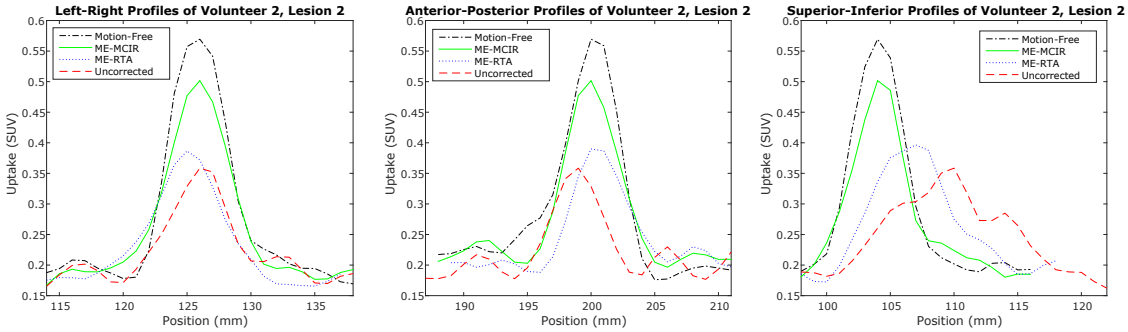

Figure 7.4: **Line profiles comparing ME-MCIR with varying techniques** The line profiles in this figure show the absolute uptake of tracer across lesion 2 in volunteer 2. This lesion was located approximately at the centre the right lobe of the liver.

7.4.1 Comparison with ME-RTA

The results for the comparison experiment are presented in several different ways. The measures of central tendency are stated in the text, with error bar or box plots used to visually compare each technique. Scatter plots are also presented. The scatter plots compare each method's performance pairwise: each plotted point is the recovery of that lesion's measurements by two methods. The line $y = x$, which represents no change, is displayed on the scatter plots to help the reader visually assess any overall trends in the comparison.

Visual Assessment

All lesions were assessed and compared qualitatively through visual assessment. Some examples of line profiles can be seen in figure 7.4 and examples of reconstructions can be seen in figure 7.5. More quantitative comparisons are drawn in the next subsection. An example of estimates of the G respiratory signals can be seen in figure 7.6.

Lesion Peak Uptake Recovery

The peak uptake of each lesion was measured using $\mathrm{SUV}_{\mathrm{peak}}$. The measured value was divided by the equivalent value in the gold-standard set by the motionless simulations to get an idea of percentage recovery of the true value. The performance

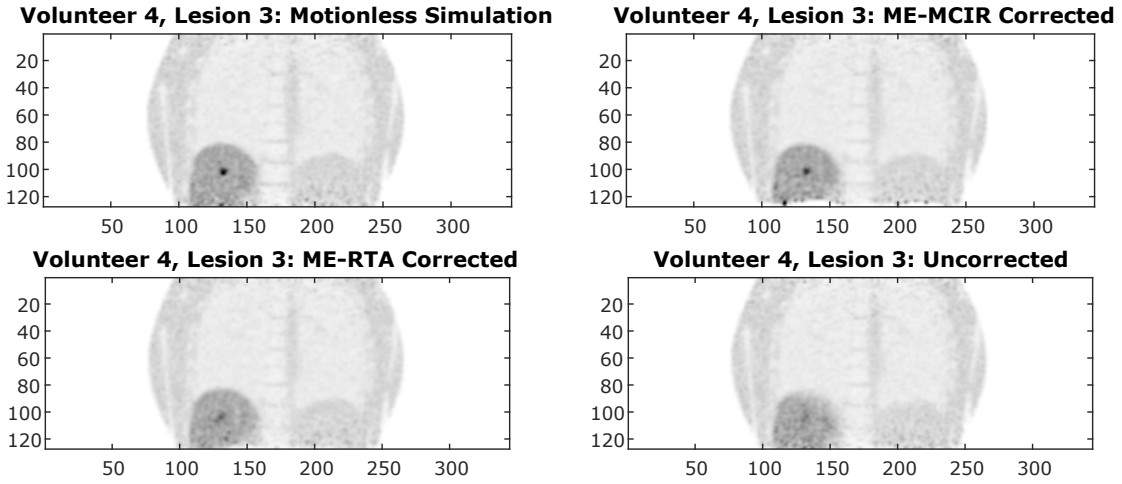

Figure 7.5: **Coronal slices of PET images reconstructed with different techniques** This figure shows coronal slices through the reconstructions of volunteer 4, with a lesion simulated in position 3. The ME-MCIR method performs relatively well at recovering the lesion shape and contrast, especially compared to the uncorrected reconstruction (bottom left) in which it is hard to see the lesion at all. Note also the loss of contrast along the diaphragm due to respiratory motion.

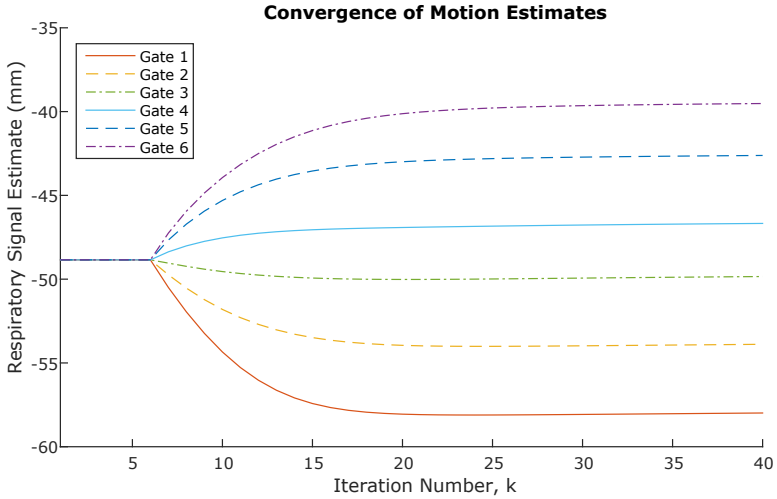

Figure 7.6: **Graph showing successive motion estimates** This graph shows successive estimates of the respiratory signal of each gate with increasing iteration number. All gates are initialised to a value in the middle of the observed range, and no motion update is performed for the first 6 iterations whilst the image sharpens, as suggested by preliminary experiments. Note that stepsize is set to the same value for all gates during the first motion update and remains constant for the remainder of the reconstruction.

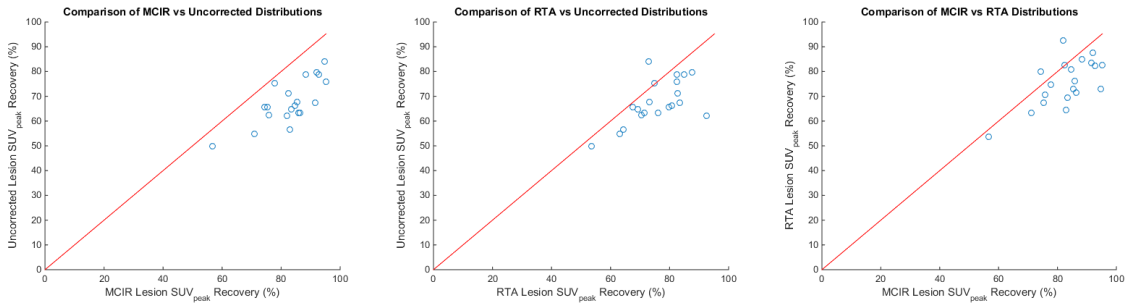

Figure 7.7: **Scatter plots: Comparative lesion uptake recovery** Each scatter plot compares two of the three methods assessed in this section. These compare the individual SUV_{peak} recoveries of all lesions simulated. The diagonal line represents $y = x$, such that any points on this line exhibited no change. In this case, the side with the most datapoints achieved better correction overall.

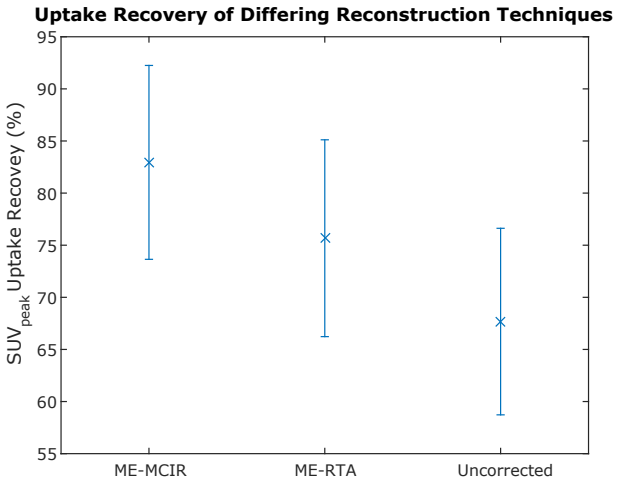

Figure 7.8: **Mean lesion uptake recovery for each method** These error bar plots compare the distributions of SUV_{peak} recovery, according to each reconstruction/correction technique.

in recovering the uptake of individual lesions is compared using scatter plots, shown in figure 7.7. Each axis corresponds to the peak uptake recovery for one of the 3 techniques (ME-MCIR, ME-RTA, or uncorrected). Points that lie on the line $y = x$ were corrected to a similar standard by the two approaches, and deviation away from this line indicates the change in performance.

The overall distribution of the recoveries of SUV_{peak} are shown on the error bar plots in figure 7.8. Error bars were used to present the data because the statistical tests described above suggested that the distribution of lesion uptake recoveries was approximately normal.

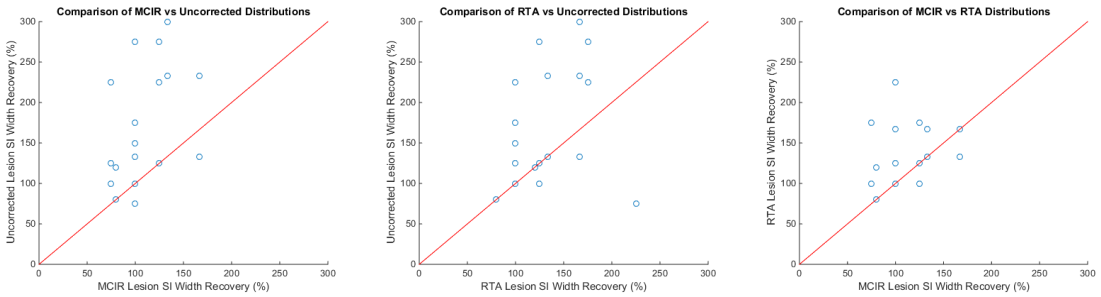

Figure 7.9: **Scatter plots: Comparative lesion size recovery** These scatter plots compare the recovery of lesion width in the superior-inferior direction for all lesions simulated, according to each reconstruction/correction technique. The diagonal line represents $y = x$, such that any points on this line exhibited no change. In this case, the side with the *fewest* datapoints performed the best.

Firstly, the uncorrected lesions only recovered the mean peak uptake to $68 \pm 8.9\%$, where the quoted error on this mean is the standard deviation. Comparatively, the mean SUV_{peak} recovery of the ME-MCIR method was observed to be $83 \pm 9.3\%$. Significance was tested using a paired two-tailed Student's t-test. The distribution of peak intensities were improved significantly compared to the uncorrected equivalents ($p = 1.9 \times 10^{-9}$). The mean uptake recovery of ME-RTA was $76 \pm 9.4\%$. This distribution was also significantly different from the uncorrected lesion uptake, with a p-value of $p = 4.7 \times 10^{-4}$.

Lesion Width Recovery

Lesion width was analysed according to the MBB method described in section 6.4.4. However, this time the analysis was performed on background-subtracted lesions, allowing for more reliable measurement of the widths without the complication of nearby image features, which can especially be a problem for lesions bordering with the diaphragm. This was done by transforming and subtracting the respective noiseless map, transformed with the motion estimates used acquired from the reconstruction. The lesions were then manually delineated using ITK-SNAP. This larger VOI formed the search space for the MBB algorithm. However, some lesions (all non–motion-corrected reconstructions) were unable to be measured in this way since they were not distinct enough from the background noise to define a sensible MBB. This occurred in 4 cases, and all have been discluded as outliers.

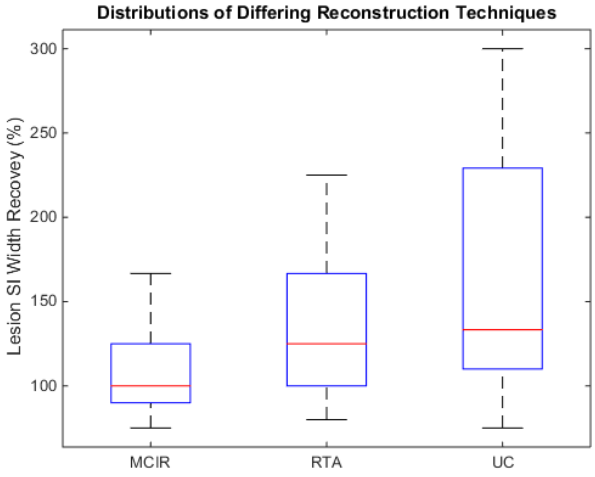

Figure 7.10: **Boxplots of size recovery** These box plots compare the distributions of lesion size recovery (in the superior-inferior direction), according to each reconstruction/correction technique.

When not using any motion correction, the superior-inferior width of the lesions was recovered to $133 \pm 119\%$. This distributions of these measurements was calculated to be non-normal, so the quoted average and associated error are the median and interquartile range. The median width recovery by the ME-MCIR method was $100 \pm 35\%$, where 7 out of 20 were recovered to the correct width. This was significantly different from the uncorrected lesion width recoveries ($p = 6.7 \times 10^{-4}$), as measured with a two-tailed paired Wilcoxon signed rank test. The width recovery of ME-RTA was $125 \pm 67\%$. However, the ME-RTA distribution of superior-inferior lesion width recoveries were not significantly different compared to the uncorrected recoveries, with $p = 0.030$.

Figure 7.9 shows comparative recovery of each individual lesion achieved with each method. The distributions for the averages above are visualised in the box and whisker plots in figure 7.10.

Lesion Position Recovery

Lesion position recovery was calculated using the position of the $\mathrm{SUV}_{\mathrm{peak}}$ measurement. The absolute displacement was calculated for each lesion. Whilst none of the sets of measurements in this case were skewed, the Shapiro-Wilk test suggested that the underlying distributions might not be normal, and so median and interquartile

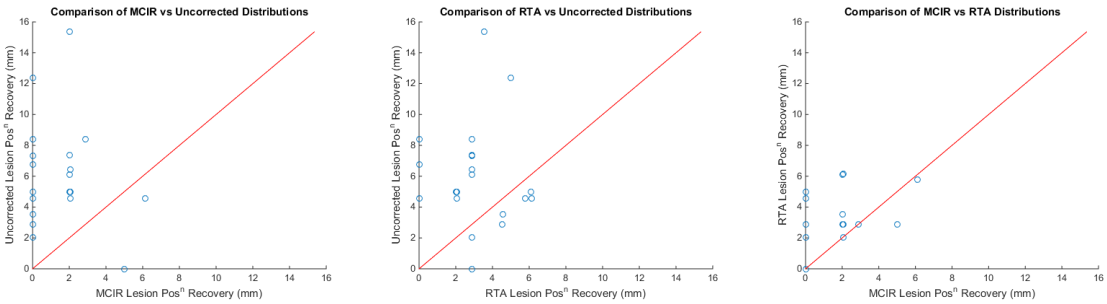

Figure 7.11: **Scatter plots: Comparative lesion position recovery** These scatter plots compare the recovery of lesion position for all lesions simulated, according to each reconstruction/correction technique. The diagonal line represents $y = x$, such that any points on this line exhibited no change. The side of the $y = x$ line containing the fewest datapoints is the better technique, on average.

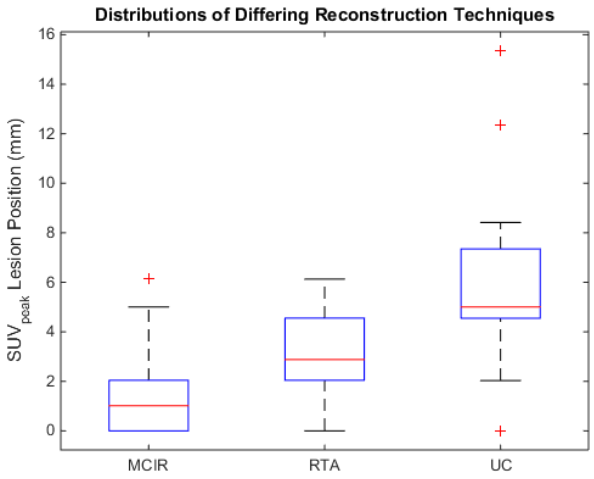

Figure 7.12: **Boxplots of lesion position recovery** These box plots compare the distributions of lesion position recovery (where position is the location of $\mathrm{SUV}_{\mathrm{peak}}$), according to each reconstruction/correction technique.

range were used.

Left uncorrected, the median lesion position shifted by 5.0 ± 2.8 mm compared to the true position in the motionless simulations. The ME-MCIR method improved the median position recovery to 1.0 ± 2.0 mm with respect to the motionless position, and ME-RTA recovered it to 2.9 ± 2.5 mm. Both ME-MCIR and ME-RTA were significantly different from the uncorrected distribution ($p = 6.0 \times 10^{-4}$ and $p = 3.8 \times 10^{-3}$, respectively). The difference between MCIR and RTA is significant, $p = 0.0038$, as decided by a Wilcoxon signed rank test.

The recovery of each lesion, compared to each other technique, is shown in the

scatter plots in figure 7.11. Figure 7.12 summarises the performance and spread of each distribution using box and whisker plots.

Reconstruction Times

Due to the nature of ME-MCIR, it took longer than the non–motion-correcting MLEM-based algorithms. This is because a large number of additional projections were required for the motion estimating part of the algorithm. For G gates, N iterations (of which, N_{wait} were without motion estimation), ME-MCIR required $(3N - 2N_{\text{wait}})G$ forward-projections and $NG + 1$ back-projections, taking average of 20.1 s and 17.5 s per projection, respectively, on an Intel Xeon X5660 processor (2.8 GHz, 6 cores, 12 threads). The next most time consuming operation was the motion transformation (with linear interpolation), at 2.9 s per operation. All numbers were calculated using MATLAB's inbuilt profiling tool. In contrast, ME-RTA required $(N + 1)G$ back-projections and NG forward-projections.

In an example 80-iteration reconstruction, ME-MCIR reconstruction took 13 hours 36 minutes, considerably slower than an equivalent ME-RTA reconstruction, which took 5 hours and 24 minutes. Altogether, projection functions and motion transformations were measured to take up 80.7% of the ME-MCIR run-time. Projection functions alone corresponded to 94.0% of the overall processing time for ME-RTA.

7.4.2 Gate Number Experiments

The same methods as those used above were employed to analyse the data in the gate number experiment. This time, however, the behaviour of the recovery methods over the number of gates were assessed. Relative success of SUV_{peak} recovery is shown in figure 7.13. This was done for both ME-MCIR reconstruction and a benchmark PET-to-PET registration based RTA method. The plot shows mean and standard error. An example of lesion appearance and position can be seen in the coronal slices displayed in figure 7.14.

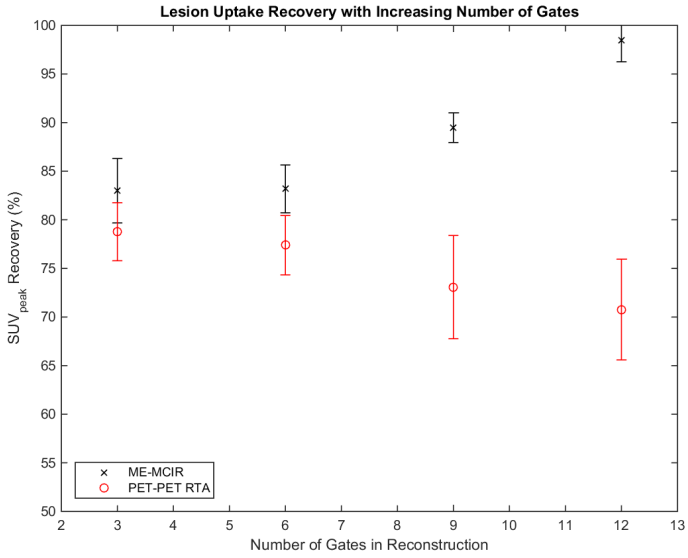

Figure 7.13: **Behaviour of ME-MCIR with varying gate number: Peak uptake recovery** This error bar plot shows the mean SUV_{peak} recovery of 5 lesions, spread over 4 volunteers, by the ME-MCIR algorithm as the number of gates in the reconstruction increases. This is compared to the PET-PET RTA, shown in red. Error bars shown are standard error of the mean.

7.5 Discussion

7.5.1 Differences to Experiments in Previous Chapter

This experiment was notably different to the original ME-RTA method in the last chapter because the lesions were deliberately placed in areas of medium-to-high displacement to avoid the systematic error of high apparent-recovery due to lesions not degrading much in the first place. This is reflected by the reduction in uncorrected SUV_{peak} compared to the previous experiment, at a *median*[2] 66% uptake recovery, compared to 78% before, which corresponds to a 12% drop, on average. Additionally, the interquartile range of the uncorrected uptake recovery distribution fell from ±18.6% to ±12.8%. The only comparable correction technique, ME-RTA, shows a similar change in median uptake recovery, which now stands at 76% instead of 87% before, corresponding to a 11% drop. This suggests that the differences in average uptake recovery are due to differences in experiment design. The +8% improvement by ME-RTA observed before is similar to the +9% improvement observed in these

[2]Whilst the mean is quoted in the results section, median is quoted here to remain comparable to the results in the previous chapter.

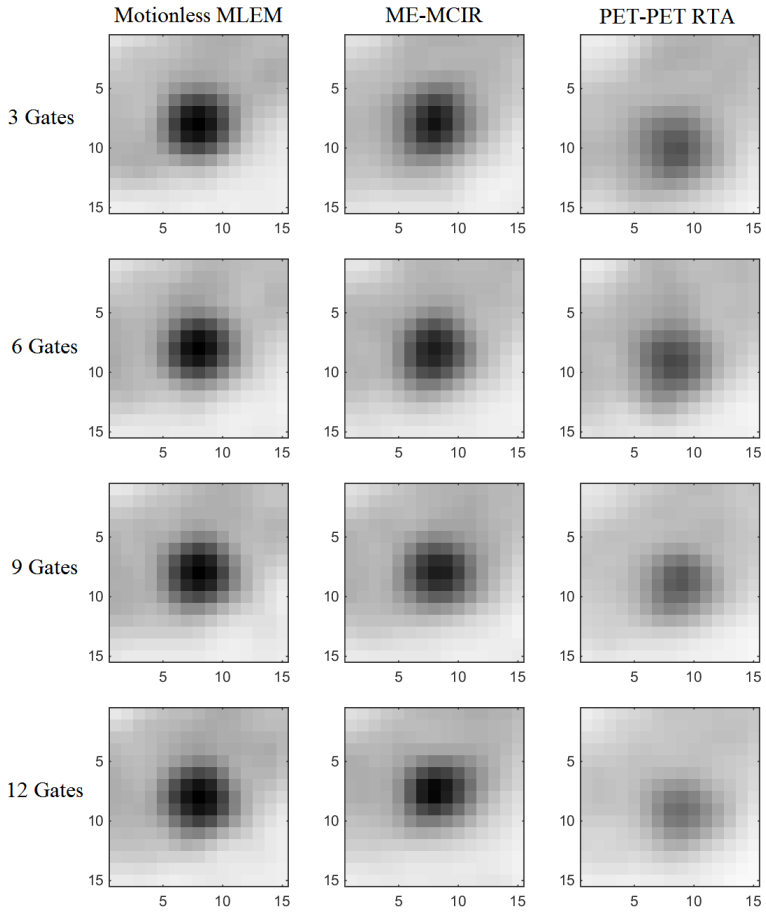

Figure 7.14: **Change in motion-corrected lesion appearance with increasing gate number** This figure displays an observed change in lesion appearance as the number of gates increases from 3 to 12. The motionless reconstruction is shown to the left, with the ME-MCIR reconstruction shown in the centre, and the PET-PET RTA shown on the right. All images are coronal slices of a myocardial lesion, shown on the same intensity scale and centred on the same position.

experiments, suggesting consistent results.

Performance Compared to ME-RTA

In general, both the ME-MCIR and ME-RTA methods were able to significantly improve on the uncorrected lesion properties, with the exception of the ME-RTA lesion SI width recovery. This could be an intrinsic failure by the ME-RTA method. However, measuring the width of motion-affected lesions can be challenging, and might also contribute to this. In future, it might be better to try fitting a Gaussian to the lesion for quantification (as done by, e.g., [152]), although it is unclear how well-defined this would be for lesions whose shape is non-spherical due to motion or poor correction. Alternatively the lesions could be simulated with much stronger contrast to minimise the effects of local background noise when defining the MBB. Care should be taken before doing this since it might artificially improve the performance of the motion estimation, which involves high-contrast edges in the image.

In terms of lesion uptake, ME-RTA provided an increase in mean uptake recovery of +8% compared to the uncorrected distribution. Motion-estimating MCIR showed considerable improvement on the uncorrected lesion measurements across the board, with +15% boost on mean $\mathrm{SUV}_{\mathrm{peak}}$ recovery, with improvement on all lesions measured. The position of $\mathrm{SUV}_{\mathrm{peak}}$ was used to mark the position of each lesion. In particular, ME-MCIR recovered a number of lesion positions to below the PET resolution (simulated at approximately 4 mm), with 10 recovered to the true position in the motionless simulations.

Some overcorrection of lesion SI width was noted for ME-MCIR, with several lesions dipping below 100% of the motionless lesion width. The exact cause of this is not clear. The motion model itself could be the cause, although the same effect was not observed for ME-RTA (for which only 1 lesion had an overcorrected width). Additionally, the motion estimates used for simulation and model formation were acquired during the same short acquisition, so it is unlikely that there is a difference in breathing pattern. Alternatively, it could be the motion estimation. For example, some gates might have had their internal variable overestimated.

ME-MCIR took approximately 2.5 times longer to compute than ME-RTA. This is somewhat expected since many joint-estimation reconstruction methods are affected by long run-times. Approaches to minimise this could be explored such as switching the gradient off once its magnitude falls below an accepted amount. Other formulations which increase the convergence speed of the motion estimation might also facilitate this method. Alternatively, other implementations of the gradient might also be possible, such as allowing a larger step every few image updates.

In general it is unsurprising that ME-MCIR outperformed ME-RTA, since it was probability-based and explicitly modelled the noise character of the data. The use of NCC in the ME-RTA technique did not do this; the ME-RTA even required smoothing prior to the NCC comparison to mitigate some of the noise problems, removing some of its sensitivity to motion. However, the motion model used in this work is relatively simplistic. This model was empirically chosen as a proof-of-principle of the technique. There is much scope for additional complexity, such as a second internal parameter, for which it would be interesting to see whether the technique introduced in this chapter would be able to achieve better or more reliable motion correction than is presented here. Some further discussion on this will be covered in the next chapter.

Performance with Gate Number

The ME-MCIR appears to cope well with increasing the number of gates. For few gates, it is not clear whether ME-MCIR performs significantly better than PET-PET registration with RTA. However, PET-PET performance clearly degrades as the number of gates increases, both in absolute lesion recovery and in reliability. On the other hand, the ME-MCIR method appears to significantly improve, with some improvement in reliability.

This could be due to a couple of reasons. Firstly, more gates means that more motion parameters were required, each of which represented fewer counts. Therefore, there is less reliance on any given parameter estimate being accurate: an error in a gate with few counts will not result in significant error in the final image. Secondly, the

amount of intra-gate motion present within the simulated PET gates reduced as the number of gates increased. Thus more motion blur could be accounted for in general, since motion within the gates would be otherwise inaccessible to this method.

The recovery of lesion width in the superior-inferior direction does not support this improvement trend. However, the error bars are large, and delineation of small, faint lesions such as those used in this experiment can be challenging. It is therefore tempting to place more trust in the SUV results. However, it would be interesting to use more datasets to improve the measurement of lesion size: errors due to the difficulties related to delineation will be at least partially reduced due to averaging, and the size of the error bars (which in this case represent standard error) would be reduced as the number of datasets increases.

The recovery of lesion position is also challenging in the context of this experiment. As we know from the comparison to ME-RTA, ME-MCIR is generally effective at recovering simulated lesion position to within a few millimetres of the motionless position. The relative variation that might occur due to a change in gate number is fairly discretised due to the location of SUV_{peak} only shifting by one or two voxels. However, it is reassuring that no significant error was observed, and it appears that increasing from 3 to 6 gates improved the lesion position estimation. The lack of error bars on the 6 gate average is likely due to the discretisation problem.

It is worth noting that there are a couple of provisos with this experiment. Since this experiment only involved relatively few datasets, any conclusions drawn remain fairly preliminary, although every care has been taken to ensure proper use of statistics. In addition, only 35 dynamic MR volumes were available in total, half of which were used to form the motion model. This left 18 for including motion into the PET simulation; at 12 gates, this is too few to suitably approximate a realistic intra-gate motion. Conversely, at 12 gates (and assuming an average navigator range of 20 mm), the motion within each gate would be approximately 2 mm at most, which is below the intrinsic resolution of a human PET scanner. Therefore, it could be argued that using this many gates makes them approximately static, since other sources of error become more significant. As such, the use of only a couple of volumes to simulate intra-gate motion might not be as unrealistic as it first seems.

Hitting the intrinsic resolution of the scanner in amplitude-gated PET studies is a limit of this approach. It would be unreasonable to expect the internal parameter to be estimated accurately if higher numbers of gates were used, although the practical need for this would be questionable. However, other formulations and applications are possible, which might allow further reduction of the SNR to test this reduced parameterisation further. For example, separate gating of inhale and exhale PET (possibly through phase-based gating) would further reduce the SNR in each gate by approximately a factor of 2. More interestingly, this method might have useful applications in dynamic PET acquisitions, since the PET frames can be very short with relatively few counts.

7.6 Conclusion

In this chapter, a maximum-likelihood–based gradient for estimating motion during PET reconstruction was presented. This method modelled the Poisson character of the noise in the data, which allowed the technique to cope well with increasing the number of respiratory gates in the reconstruction. This meant less intra-gate motion, which a preliminary study suggests improved the ability to motion correct as the number of gates increased. This robustness to noise was also facilitated by a drastic reduction in the number of motion parameters being estimated, at 1 parameter per PET gate. This was achieved by employing an indirect-correspondence parameterised MR-derived motion model, providing a strong constraint on the possible motion estimates that could be formed. Other methods in the literature derived motion estimates purely from MR or CT data with no further optimisation, whereas others calculated a relatively large number of motion parameters from the PET data itself during reconstruction. In this sense, the method outlined here can be considered a half-way house.

However, the method presented is still quite simplistic. It has only been tested on emission-only simulated data, where other sources of error (such as attenuation) have been completely ignored. This was also a major limitation of the ME-RTA approach, to which ME-MCIR was compared. In the next chapter, including attenuation into

the problem is introduced, which has a non-trivial effect on the ability to use ME-MCIR. This is the first step towards including additional error terms, which may eventually lead to trying ME-MCIR on real data.

Chapter 8

Incorporating Attenuation into Motion Estimation

8.1 Introduction

Attenuation is a leading cause of error in PET. Historically, this was addressed using transmission scans, which used a 511 keV photon source to find the difference in received signal between a scan with an empty FOV and one with the subject present. This had the advantage of being subject to the same motion effects observed in PET emission scans. The advent of PET-CT became popular, and, although limited by the measurement of attenuation using the lower-energy photons, still formed a silver standard for estimation of attenuation effects. The move to PET-MR has yet to match the efficacy of PET-CT in attenuation estimation due to the nature of the MR signal [183] although there are efforts underway to do so (see, for example, [13]).

A problem in PET-CT is that, compared to the respiratory period, the CT is not generally fast enough to capture snapshot static images. Methods to address this include acquisition of gated 2D slices, helical scans, or cine-CT. These can still be prone to artefacts, however, and so breath-hold images are usually clinically preferred. Either of these methods can lead to artefacts in the PET, either by truly erroneous areas of perceived attenuation (such as floating liver dome artefacts in the CT), or by mismatched tissues (such as those observed between a breath-hold CT

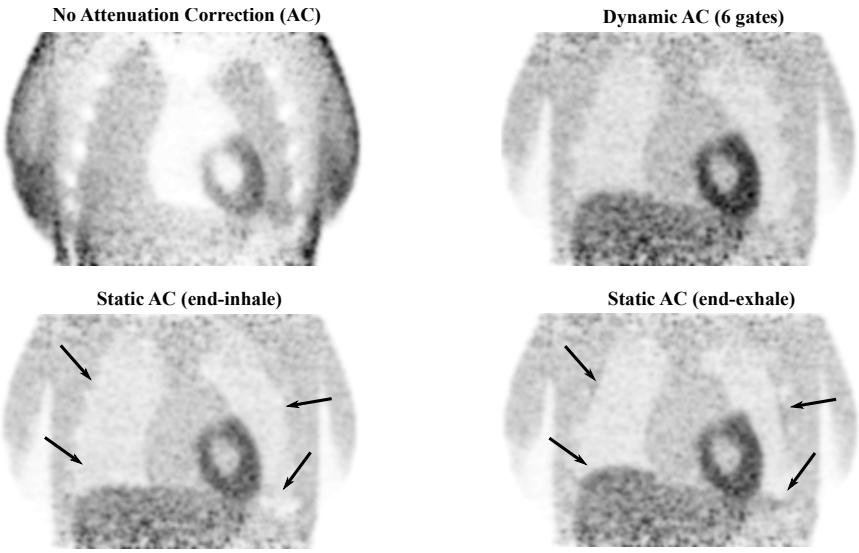

Figure 8.1: **Effects of μ-map position on uptake estimation** This compares a simulation of PET data with no attenuation correction (top left) against attenuation-corrected images. The end result is heavily dependent on the respiratory position of the μ-map, where the motion-blur of the underlying data only being apparent if a comparatively-blurred μ-map is used (top right). Using static maps (in a single respiratory position, bottom row) can change the apparent extent of motion on the image, depending on the underlying differences in contrast. Arrows highlight the differences due to the use of end-inhale or end-exhale μ-maps.

and a free-breathing PET scan) [184, 185]. In general, a CT scan cannot match the attenuation distribution in a PET scan without having the same motion character (i.e., continuously blurred), as illustrated in figure 8.1. An approximation to this can be achieved by using a set of respiratory-gated CT images, although this increases the radiation exposure to the subject [169]. This is less of a problem in motion-corrected PET, provided that the motion correction is reliable and the CT image is in the same position as that of the corrected PET.

Few of the joint motion and activity estimation approaches include attenuation, since incorporation into the objective functions can be difficult. This was recently performed by [97], who used B-spline motion fields to estimate motion in penalised likelihood, with two quadratic penalty terms.

Similar problems were met trying to introduce attenuation into the method introduced in this thesis. The previous chapter introduced an approach for estimating respiratory motion position during MCIR, using an analytical gradient based on the reduced parameterisation of a motion model. However, this method relied upon

the prompt sinogram (i.e. the data that would be retrieved from a PET scan) containing an emission-only signal, rather than being affected by complications such as attenuation, scatter, or randoms. In a real scan, this is not possible. Preliminary experiments were performed to include these effects in the image reconstruction, but naïvely used the aforementioned gradient to estimate the motion positions. This resulted in the motion estimation algorithm to produce meaningless results. Instead, the data model needs to be improved by including other sources of error in the data. First and foremost would be attenuation. This is the focus of the final methodology chapter in this thesis.

Generally speaking, such a method should only assume that one μ-map is available: having one for each gate would circumvent the need for this method, since the maps could be used to generate gatewise motion estimates. The availability of a single attenuation map is more realistic in a clinical scenario: current methods of attenuation estimation in PET-MR – such as UTE or Dixon scans – are relatively slow, making multiple acquisitions impractical. This is discussed in more detail in section 3.2.4. Finally, should one decide to extend this method to arbitrarily high numbers of gates (or perhaps even frames in a dynamic scan), having a sufficient number of pre-generated μ-maps might not be viable at all.

8.2 Background

As discussed in chapter 7, the analytical motion gradient is based on the model of the mean of the data. This has the form shown in (7.2), repeated here for convenience:

$$\boldsymbol{q} = \boldsymbol{AMf},$$

where $\boldsymbol{q}$ is the model of the mean of the data, as a sinogram, $\boldsymbol{A}$ and $\boldsymbol{M}$ are matrix operators describing the transformations imposed by the geometry and motion, respectively, and $\boldsymbol{f}$ is an estimate of the activity distribution.

Representing this data model with continuous functions and operators facilitated the derivation of an analytical gradient of the Poisson log-likelihood objective function

with respect to a scalar internal variable, s. This parameter allowed a drastic reduction in the number of additional parameters that needed to be estimated to obtain motion estimates during PET image reconstruction. This is because it formed part of an IC parameterised motion model, built from dynamic MR volumes, forming a strong constraint on the positions the voxels could take over the course of the respiratory cycle.

The internal variable was estimated for each PET gate, and subsequently updated during the second half of each iteration of the image reconstruction. The gradient for updating the motion parameters is presented in (7.22). The full form of the emission-only gradient is

$$\frac{\mathrm{d}L(s|\boldsymbol{m}, \boldsymbol{f}^{(k+1)})}{\mathrm{d}s} = \sum_{i=1}^{I} \left(\frac{m_i}{[\mathcal{R}\{g(\vec{r}, s)\}]_i} - 1 \right) \left[-\mathcal{R} \left\{ \sum_{j=1}^{J} \vec{\phi}'_j(s) \cdot \vec{\nabla}_{\vec{v}_j} b(\vec{v}_j) f_j \right\} \right]_i . \tag{8.1}$$

A brief recap of the terms in this equation is as follows: The derivation involved a continuous formulation of the data model, which involved the X-ray transform operator, $\mathcal{R}$ and a non-rigid transformation operator, solely dependent on s via a motion model, $\vec{\phi}$. The motion transformation acted upon the (estimated) activity distribution f to produce a transformed distribution g, constructed from basis functions, b, warped to a new co-ordinate grid, $\vec{\boldsymbol{v}}$. The term on the left of the right-hand side controls the gradient during convergence. The term on the right lends weight to specific aspects of the image, which involves the derivative of the motion model, $\vec{\phi}'$. This contains a Cartesian dot product of the gradient operator with the derivative of the motion model. It is at this point that the gradient was discretised, represented by the square brackets.

8.2.1 Improving the Data Model

As discussed in the introduction, preliminary experiments found that including attenuation in the image update but not the motion update produced nonsense results. More scrutiny was paid to the original derivation of the ME-MCIR update equation to see whether an answer could be found in the underlying data model.

Equation 7.5 was improved by modifying the model of the data such that it would explicitly contain an attenuation term. For a general motion field, $\Delta\vec{r}$, this is

$$q(\vec{u}, \Delta\vec{r}) = a(\vec{u}, \Delta\vec{r})\mathcal{RT}\left\{f(\vec{r}), \Delta\vec{r}\right\}. \tag{8.2}$$

Here, a is the continuous equivalent to the attenuation factor (AF) sinogram discussed in section 2.1.4. It has been modified to include the transformation due to motion, as described by displacement field $\Delta\vec{r}$. Equation 2.2 can be used to show that

$$a(\vec{u}, \Delta\vec{r}) = \exp\left(-\mathcal{RT}\left\{\mu(\vec{r}), \Delta\vec{r}\right\}\right) \tag{8.3}$$

which relies on the spatial distribution of attenuation coefficients, $\mu(\vec{r})$.

As before, using the motion model, $\vec{\phi}$, to constrain $\Delta\vec{r}$ to a single internal variable, s, we arrive at the full expression for the attenuation-included data model,

$$q(\vec{u}, s) = \exp\left(-\mathcal{RT}\left\{\mu(\vec{r}), \vec{\phi}(s)\right\}\right)\mathcal{RT}\left\{f(\vec{r}), \vec{\phi}(s)\right\}. \tag{8.4}$$

The AF sinogram, a, depends upon the same motion transformation as the emission sinogram. This suggests that it might also contribute to updates of the internal variable s when derived using the log-likelihood method described in the previous chapter.

8.2.2 Derivative of the Improved Log-Likelihood

Assuming the derivative of the Poisson log-likelihood takes the same form as (7.14), we must recalculate the derivative of q with respect to s. For clarity, let (8.4) be split into two terms such that $q(\vec{u}, s) = a(\vec{u}, s)l(\vec{u}, s)$. In a continuous basis,

$$\frac{\partial q}{\partial s} = \frac{\partial a}{\partial s}l + a\frac{\partial l}{\partial s} \tag{8.5}$$

where the second term on the right is a multiplied by the gradient already derived in the previous chapter (refer to (7.21)). Notice that since a contains an exponential term, it will persist in the derivative. For clarity, if we make the substitutions

$\nu(\vec{r}, s) = \mathcal{T}(s)\{\mu(\vec{r})\}$ and $\alpha(\vec{u}) = -\mathcal{R}\,\nu(\vec{r}, s)$, then

$$\frac{\partial a}{\partial s} = \frac{\partial \alpha}{\partial s} \exp \alpha \equiv -\mathcal{R}\left\{\frac{\partial \nu(\vec{r}, s)}{\partial s}\right\} \exp\left(-\mathcal{R}\,\nu(\vec{r}, s)\right), \tag{8.6}$$

where the exponential term is simply a, as before.

By analogy, we have computed the derivative term previously, in (7.21). Expanding in terms of image basis functions,

$$\frac{\partial\,\nu(\vec{r}, s)}{\partial s} = -\sum_{j}^{J} \vec{\phi}'_j(s) \cdot \vec{\nabla}_{\vec{v}_j} b(\vec{v}_j)\,\mu_j, \tag{8.7}$$

where $\vec{v}_j = \vec{r} - \vec{r}_j - \vec{\phi}_j(s)$, as defined in section 7.2.3.

Given that we have the data model $q = al$, where a is the attenuation coefficient sinogram, and l is the emission sinogram in the previous chapter, we find that (7.22) becomes

$$\frac{\mathrm{d}L(s|\boldsymbol{m}, \boldsymbol{f}^{(k)})}{\mathrm{d}s} = \sum_{i=1}^{I} \left(\frac{m_i}{[al]_i} - 1\right) [X + Y]_i \tag{8.8}$$

where

$$X = -a(\vec{u})\mathcal{R}\left\{\sum_{j=1}^{J} \vec{\phi}'_j(s) \cdot \vec{\nabla}_{\vec{v}_j} b(\vec{v}_j) f_j\right\} \tag{8.9}$$

and

$$Y = a(\vec{u})\mathcal{R}\left\{\sum_{j}^{J} \vec{\phi}'_j(s) \cdot \vec{\nabla}_{\vec{v}_j} b(\vec{v}_j)\,\mu_j\right\} l(\vec{k}). \tag{8.10}$$

The X term is equivalent to the gradient weights seen in the emission-only case, but multiplied by the AF sinogram. The second term, Y, is new. It involves the data model (i.e. al), multiplied by the sinogram of weights. However, this time, the weights sinogram is derived from the μ-map, rather than the emission image. As we would expect, Y vanishes in the absence of attenuating material (i.e. $\boldsymbol{\mu} = 0$), returning us to the emission-only gradient, (7.22). These two changes, the attenuation of the emission term and the emergence of a new counter-term for the μ-map, were found to once again provide sensible motion estimates in preliminary experiments including attenuation.

Discretisation of (8.8) requires no extra machinery in addition to what has been

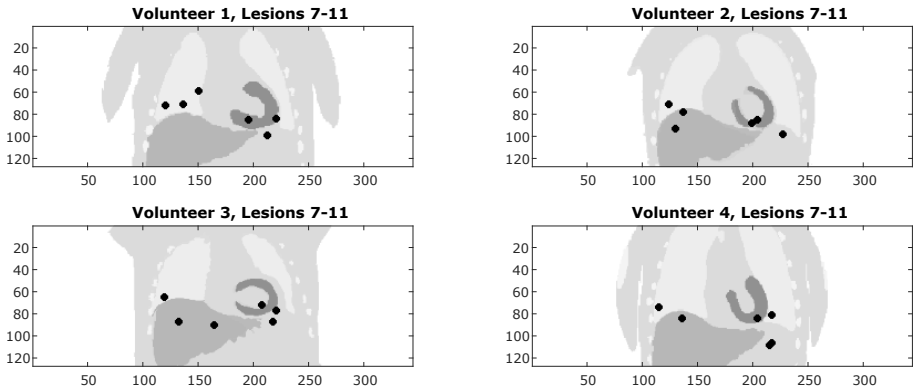

Figure 8.2: **Additional lesion locations projected onto coronal slice** This figure shows where 5 additional lesions were positioned for each volunteer, projected onto a coronal slice. These are in addition to those shown in figure 7.2 to improve the statistical power of the data in this chapter. These positions were selected due to the higher than average displacement observed in the motion fields as before. Note that the tissue type of the lesion in these images might not reflect the true tissue type in which it is placed.

developed already. As before, we use the voxel representation, which allows us to represent $\vec{\phi}' \cdot \vec{\nabla}$ by a weighted matrix of central differences, $\boldsymbol{D}_{\mathrm{W}}$. Thus,

$$\frac{\mathrm{d}L}{\mathrm{d}s} = [\boldsymbol{a} \circ \boldsymbol{A}\boldsymbol{D}_{\mathrm{W}}\boldsymbol{M}\boldsymbol{\mu} \circ \boldsymbol{A}\boldsymbol{M}\boldsymbol{f} - \boldsymbol{a} \circ \boldsymbol{A}\boldsymbol{D}_{\mathrm{W}}\boldsymbol{M}\boldsymbol{f}]^{\mathrm{T}} \left(\frac{\mathbf{m}}{\boldsymbol{a} \circ \boldsymbol{A}\boldsymbol{M}\boldsymbol{f}} - \mathbf{1} \right), \tag{8.11}$$

where $\boldsymbol{A}$ and $\boldsymbol{M}$ are the projection and motion matrices, respectively, $\boldsymbol{a}$ is the AF sinogram, $\boldsymbol{m}$ the data sinogram, and $\boldsymbol{f}$ and $\boldsymbol{\mu}$ are the emission and μ-map images, respectively. Note that $\circ$ is used to denote the Hadamard product – In this context, this corresponds to elementwise multiplication of two vectors of the same length. This could otherwise be replaced by a matrix multiplication with $\mathrm{diag}(\boldsymbol{a})$.

8.3 Methods and Materials

8.3.1 Simulation Details

As in the previous two chapters, the reconstruction algorithm proposed in this chapter is tested using MR-based PET simulations. The same MR datasets were used as before, with 4 volunteers, each with 35 dynamic volumes and an accompany-

ing UTE dataset. The segmented PET maps generated from the UTE data were designed to be relevant to both tracer uptake (with varying intensity in the liver and myocardium, for example), and attenuation properties (in which case the liver and myocardium are approximately the same, at $\mu = 0.096\,\mathrm{cm}^{-1}$). As such, the PET maps were used to generate μ-maps for the purposes of attenuation simulation and attenuation correction. In the experiments in this chapter, a further 5 12 mm-diameter lesions were simulated for each volunteer for better statistics, resulting in a total of 11 lesions for each volunteer in a range of positions and tissues, as shown in figure 8.2. As in the previous chapter, lesions were simulated with 3:1 background contrast on the emission map. Note that lesions were simulated to be invisible on the μ-maps.

The attenuation maps used for simulation and correction were treated differently. In the simulation of PET data, the position of each emission map and attenuation map were identical, as would be expected in a real data scenario. This includes the modelling of intra-gate motion as discussed in the previous chapter, section 7.3.2. As a result, the attenuation maps used for data creation were blurred to precisely match the FDG maps. Conversely, to model the way μ-maps are acquired in PET-CT (or even, perhaps, PET-MR), the μ-maps used for attenuation correction were *not* blurred by motion. The correction μ-maps therefore had sharp feature edges.

In general, μ-maps are not used directly for simulation (or correction) of attenuation; instead an AF sinogram must be formed. This was also the case in this experiment. The AF sinograms were formed by applying the exponential term in the attenuation law in (2.3). The resulting AF sinogram has values which range between 0 and 1, representing the average proportion of 511 keV photons that are attenuated along that LOR. As detailed above, the data creation μ-maps were treated differently to those used for correction. The method of generating the AF sinograms from either was the same.

8.3.2 Reconstruction

The reconstruction pathway is similar to that used in the previous chapter: image and motion parameter estimation are interleaved, such that each is separately updated during each iterative update. Each image update was performed according to the MCIR update equation, during which the respiratory motion was assumed unchanging. This is similar to (7.30), although this time attenuation must be included [140, 186] as follows:

$$\boldsymbol{f}^{(k+1)} = \frac{\boldsymbol{f}^{(k)}}{\sum_{g=1}^{G} \boldsymbol{a}_g \circ \boldsymbol{M}_g^{\mathrm{T}} \boldsymbol{A}^{\mathrm{T}} \mathbf{1}} \sum_{g=1}^{G} \boldsymbol{M}_g^{\mathrm{T}} \boldsymbol{A}^{\mathrm{T}} \frac{\boldsymbol{m}_g}{\boldsymbol{A}\boldsymbol{M}_g \boldsymbol{f}^{(k)}} \ . \tag{8.12}$$

The internal variables for driving the motion model were updated gatewise using gradient ascent as before. This requires discretisation of the gradient in (8.11). This results in

$$\begin{aligned} s_g^{(k+1)} = s_g^{(k)} + \beta^{(k)} \left(\frac{\boldsymbol{m}_g}{\boldsymbol{a}_g \circ \boldsymbol{A}\boldsymbol{M}_g \boldsymbol{f}^{(k+1)}} - \mathbf{1} \right)^{\mathrm{T}} \big(& \boldsymbol{a}_g \circ \boldsymbol{A}\boldsymbol{M}_g \boldsymbol{f}^{(k+1)} \circ \boldsymbol{A}\boldsymbol{D}_{\mathrm{W},g} \boldsymbol{M}_g \boldsymbol{\mu} \\ & - \boldsymbol{a}_g \circ \boldsymbol{A}\boldsymbol{D}_{\mathrm{W},g} \boldsymbol{M}_g \boldsymbol{f}^{(k+1)} \big) \end{aligned} \tag{8.13}$$

where

$$\boldsymbol{D}_{\mathrm{W},g}^{(k)} \equiv \vec{\phi}' \left(s_g^{(k)} \right) \cdot \vec{\boldsymbol{D}}. \tag{8.14}$$

Note that $\boldsymbol{a}$, $\boldsymbol{M}$, and $\boldsymbol{D}_{\mathrm{W}}$ depend directly on the parameter estimate, $s_g^{(k)}$, and that the gradient update uses the updated image, which is updated just before it.

These two updates were performed in an interleaved fashion. As explained in the previous chapter, only the image was updated for the first 6 iterations, after which the image and the motion estimates were updated once each per iteration for a further 34 iterations, such that there were 40 iterations in total. The stepsize set such that the mean of first set of absolute gradient estimates was normalised to one millimetre per iteration, to set the scale of the gradient. A VOI was used, similar to before, placed over the right lung and liver, since preliminary experiments suggested that this improved motion estimation. Finally, the image starting estimate was equal to $\mathbf{1}$, and the motion estimates were all initialised to be in the centre of the

range of values observed during the formation of the motion model. This can be seen in figure 8.5.

8.4 Results

The results will be described in this section. First the qualitative aspects of the attenuation-incorporated scheme will be discussed, followed by quantitative results. These will compare attenuation-incorporated ME-MCIR with the emission-only simulation reconstructions. The increased number of lesions also allows a regional analysis due to improved statistics.

8.4.1 Visual Inspection

Some notable examples of the lesions are shown in figure 8.3. The top example shows the reconstructed images for volunteer 4, lesion 3. This is an example of one of the better reconstructions, with good agreement between the emission-only uptake distribution and the attenuation-included uptake distribution. The bottom example, volunteer 2, lesion 4, is relatively worse, with some notable discrepancies in the uptake distributions. In particular, the latter example has a notable attenuation artefact: a hypointense region above the liver, corresponding to a mismatch between the tissue type in the emission image estimate and the μ-map. The respective internal variable estimates are presented in figure 8.5.

Attenuation Effects

Attenuation can be a secondary source of motion-based error, as illustrated in 8.1. As such, the existence of attenuation artefacts in images can be a good indication that there is, in general, an incorrect motion estimate. The appearance of artefacts in the motion corrected images was variable, with all notable artefacts appearing as a hypointensity above the left hemidiaphragm, although one was observed between the corner between the liver dome and the spine. An example of this can be seen

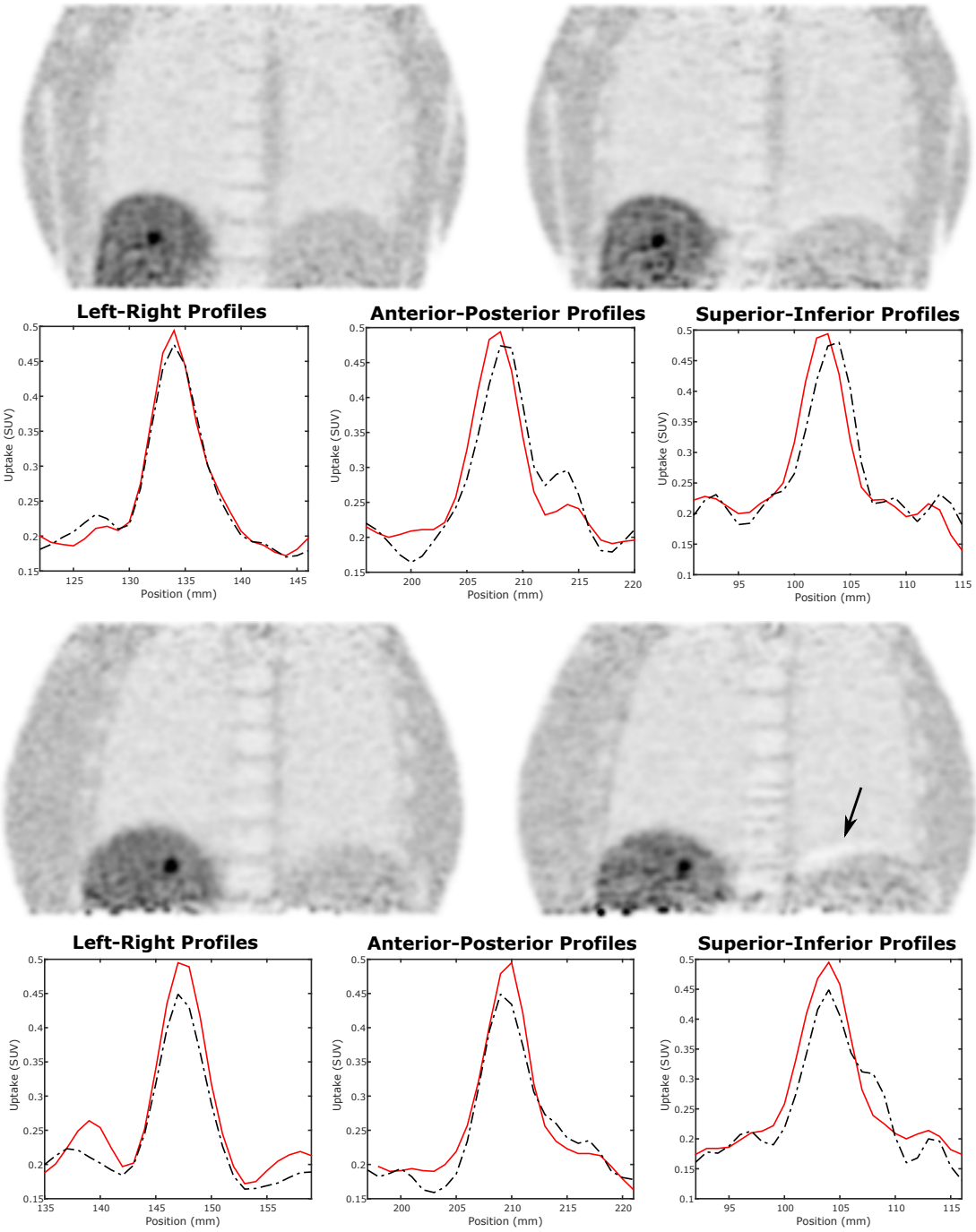

Figure 8.3: **Visual comparison of motion-corrected emission-only data vs. attenuation-included data** This figure shows some examples of images reconstructed from the emission-only dataset (left) with attenuation-included data reconstructions (right), reconstructed using the respective versions of ME-MCIR for volunteer 4, lesion 3 (top) and volunteer 2, lesion 4 (bottom). Below each example are the lesion profiles in each direction, where the red solid line denotes the uptake of the emission-only lesion and the dashed black line represents the attenuation-included lesion. An arrow can be seen in the lower-right image, highlighting an attenuation artefact over the left hemidiaphragm.

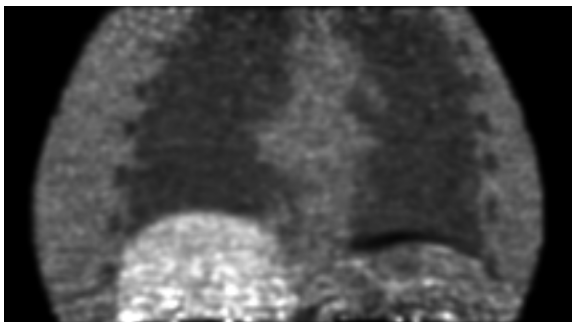

Figure 8.4: **A high-contrast image of the worst observed attenuation artefact** Here a hypointensity can be observed above the left hemi-diaphragm. In terms of extent and difference in intensity, this was the worst observed attenuation artefact.

in figure 8.3, but the worst example observed is presented in high contrast (for visibility) in figure 8.4. Some discrepancies in the spinal uptake can also be seen in figure 8.3 in the lower right coronal slice. Apart from one case mentioned above, no artefacts were observed above the right hemidiaphragm.

Observations

On inspection of the reconstruction data, it was observed that some datasets were able to estimate motion as well as in the emission-only case, whereas others were not. In the emission-only reconstruction, estimates were fairly uniformly distributed and all converged at a roughly similar rate. The uniform distribution is expected since the gates were equally-sized intervals of the respiratory amplitude. All respiratory signals were also observed to stay in the correct order. However, when attenuation was included, some datasets became irregular. In worse-case scenarios, such as that shown on the centre right of figure 8.5, some gates converge to the same point, or even crossed over such that the ordering was no longer as it should be. There is no clear indication in the images to hint at what caused this behaviour, although some possibilities will be presented in the discussion of this chapter.

8.4.2 Quantification

A total of 44 lesions were simulated, but 7 of these were removed from quantitative analysis as outliers. This is because, despite background subtraction, it was too

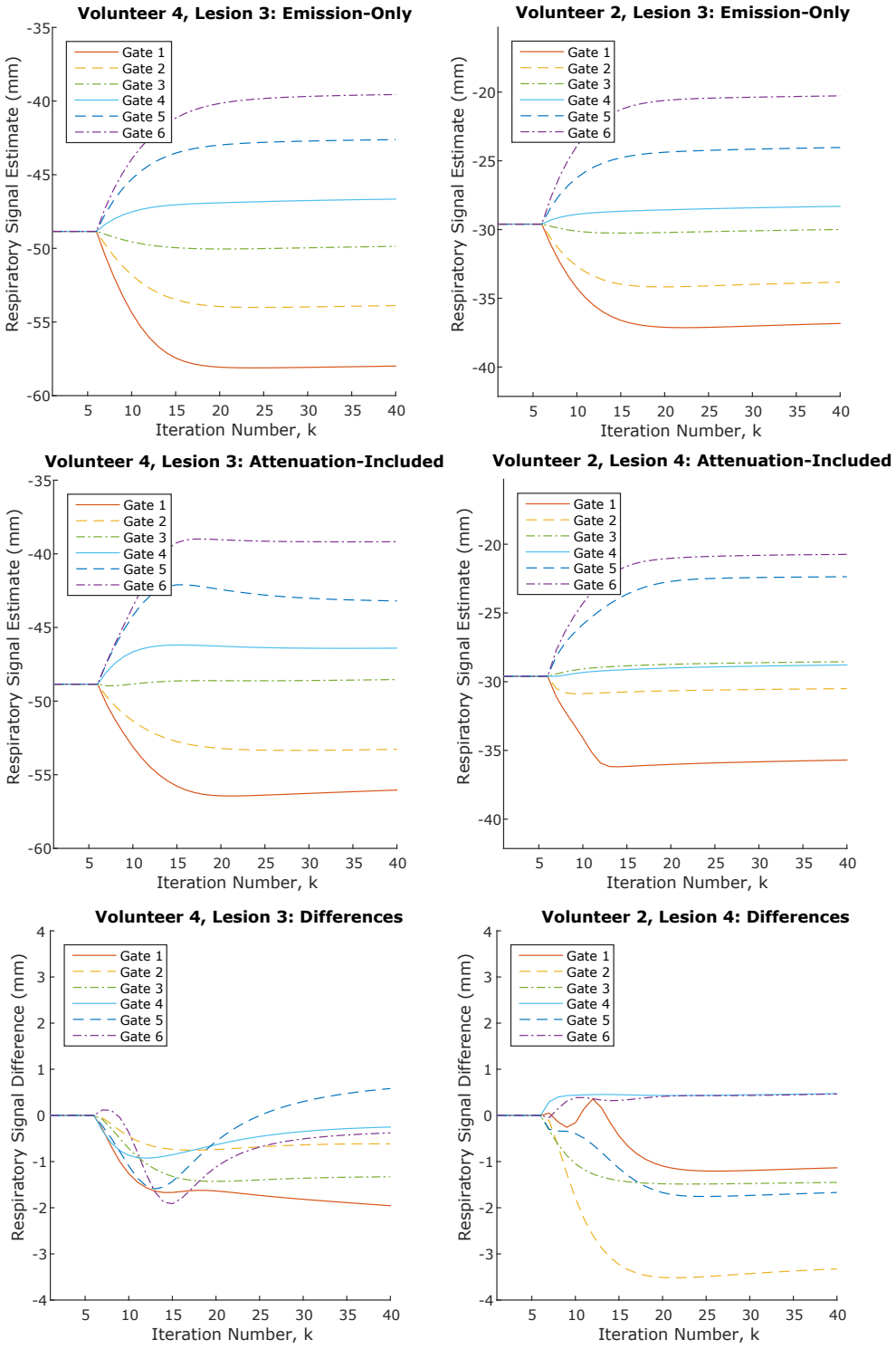

Figure 8.5: **Differences in respiratory signal convergence** This figure displays the different paths of convergence for the respiratory signal estimates in two cases: lesion 3 in volunteer 4, and lesion 3 in volunteer 2. The estimates for the respiratory signal in each of 6 gates at each iterative update can be seen. Performance varied when attenuation was included into the estimator: the emission-only gradients are stable and fairly evenly spread, as can be seen in the top row. The addition of attenuation barely affected some reconstructions, such as on the left. The drop in performance is more pronounced in some reconstructions (e.g. the right column). Gate 2 is the most strongly affected, as shown in the difference between emission-only and attenuation-included, lower right. Note that gate 6 is the reference position that the rest were corrected to, corresponding to end-expiration. Respiratory signal estimates are negative because the navigator signal is relative to the centre of the image.

difficult to detect the lesions from noise, leading to outliers in the data. Therefore, all of the following quantified analysis uses a total of 37 lesions. All of these lesions were lung lesions.

The main quantitative aim in this chapter is to assess any differences between ME-MCIR in the emission-only case in the previous chapter, and the attenuation-included case presented here. Ideally, the algorithm should perform similarly to the previous experiments for which the data was emission-only. Good correlation in performance which would suggest good motion and attenuation correction collectively.

To test this, the relative error for each of the usual metrics (peak lesion uptake, lesion width, and lesion position) were calculated and assessed. A regional comparison was also performed to see whether any tissues are disproportionately affected by this.

Peak Lesion Uptake

Peak lesion uptake recovery, calculated using $\mathrm{SUV}_{\mathrm{peak}}$, has been a staple quantifier in this thesis, and this chapter is no different. This time, however, the relative error compared to the emission-only simulations is compared. If x_{noA} is the peak uptake of a lesion in the reconstructed image of the emission-only simulations using ME-MCIR, and x_{wA} is the equivalent uptake simulated to include attenuation and reconstructed with the attenuation-incorporated ME-MCIR method, the relative error ϵ is defined as

$$\epsilon = 100 \cdot \frac{x_{\mathrm{wA}} - x_{\mathrm{noA}}}{x_{\mathrm{noA}}}. \tag{8.15}$$

This normalises the absolute uptakes of different lesions and allows comparison of both techniques using all lesions.

The skew of the resulting distribution of relative errors, ϵ_n, of N lesions was assessed using (7.34) in addition to calculating the likelihood of $\{\epsilon_n\}$ being samples of a normal distribution using the Shapiro-Wilk test. It was found that there was no significant skew and that the data is likely to be normal, since the null hypothesis was unable to be rejected ($p = 0.216$). The mean difference and associated standard deviation were $-1.6 \pm 5.9\%$, suggesting a small reduction of peak uptake estimation

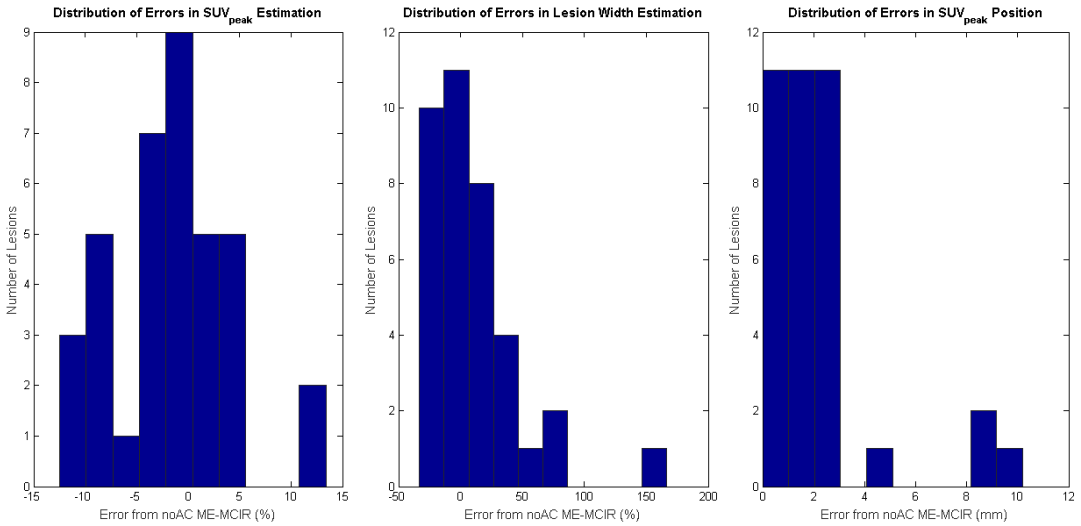

Figure 8.6: **Distributions of relative errors from emission-only measurements** These histograms show the distributions of of error for different lesion characterisation measurements. Left is the relative error of peak lesion uptake, whereas centre is the relative error in estimation of lesion width in the SI direction. These are both in units of percent. The histogram on the right is the error in lesion position in millimetres. Notice that the central and right distributions are skewed, and thus unlikely to be sampled from a normal distribution.

on average. A two-tailed Student's t-test suggested that this difference is not statistically significant, with $p = 0.0512$. The distribution of these measurements can be seen on the left hand side of figure 8.6.

Superior-Inferior Lesion Width

Lesion width in the superior-inferior direction was measured using the MBB technique introduced in section 6.4.4. These were compared between the emission-only benchmark and the attenuation-included reconstructions using relative error, as defined in (8.15).

As can be seen in the central histogram in figure 8.6, there is a noticeable skew to the error distribution, suggesting the distribution is not normal. This was supported by the Shapiro-Wilk test, which found a significant difference between these data and the normal distribution ($p < 0.001$). Therefore, the median must be used, which suggested an average difference of $0 \pm 39.9\%$ in width compared to the emission-only data. A Wilcoxon signed rank test implied no significant difference between the two

datasets ($p = 0.393$).

Position Recovery

Finally, the position of SUV_{peak} was calculated. In this case, the absolute, rather than relative, difference between each respective lesion measurement was calculated. Therefore, these values are in millimetres.

There is a significant skew to this dataset, seen on the right of figure 8.6. This dataset is significantly different to a normal distribution ($p < 0.001$), with a median difference of 2.0 ± 2.9 mm. In contrast to the previous two datasets, these distributions were found to be statistically distinct by a Wilcoxon signed rank test ($p < 0.001$).

8.4.3 Regional Analysis of Uptake Measurement

An advantage of more lesions per volunteer is the improved statistics and increased flexibility for creating subsets of data. In this experiment, subsets were generated according to where lesions were placed. Since lesions in this dataset were generated with a constant contrast (rather than peak uptake), there were tissue-type–related differences in lesion uptake estimation.

Lesions were simulated in 4 tissue types: liver, lung, heart (myocardium), and soft tissue (generally placed below the left lung). The absolute measurements of peak uptake were compared directly, which are summarised in table 8.1. The distributions of the individual values measured by each technique can be seen in the scatter plots in figure 8.7.

8.5 Discussion

In this section, the performance due to including attenuation effects into both the reconstruction problem and the data model will be discussed. Looking into the

Tissue Type	N	Mean Difference	Emission-Only	Atten^{n}-Included
Liver	17	0.006 ± 0.024 $(p = 0.31)$	0.43 ± 0.04	0.42 ± 0.04
Lung	10	0.002 ± 0.005 $(p = 0.27)$	0.10 ± 0.01	0.10 ± 0.01
Myocardium	4	0.028 ± 0.032 $(p = 0.18)$	0.61 ± 0.16	0.58 ± 0.16
Soft Tissue	6	0.000 ± 0.015 $(p = 0.94)$	0.21 ± 0.02	0.21 ± 0.02

Table 8.1: **Regional comparison of lesion peak uptake** This table summarises the measurements for all lesions, separated according to the tissue into which they were segmented. The second column displays the number of lesions in that category. The next column displays the overall difference and the p-value found by assessing whether any significant difference could be identified using a 2-tailed t-test. No distributions were significantly different. Note that all quoted values are mean $\mathrm{SUV}_{\mathrm{peak}}$ and standard deviation.

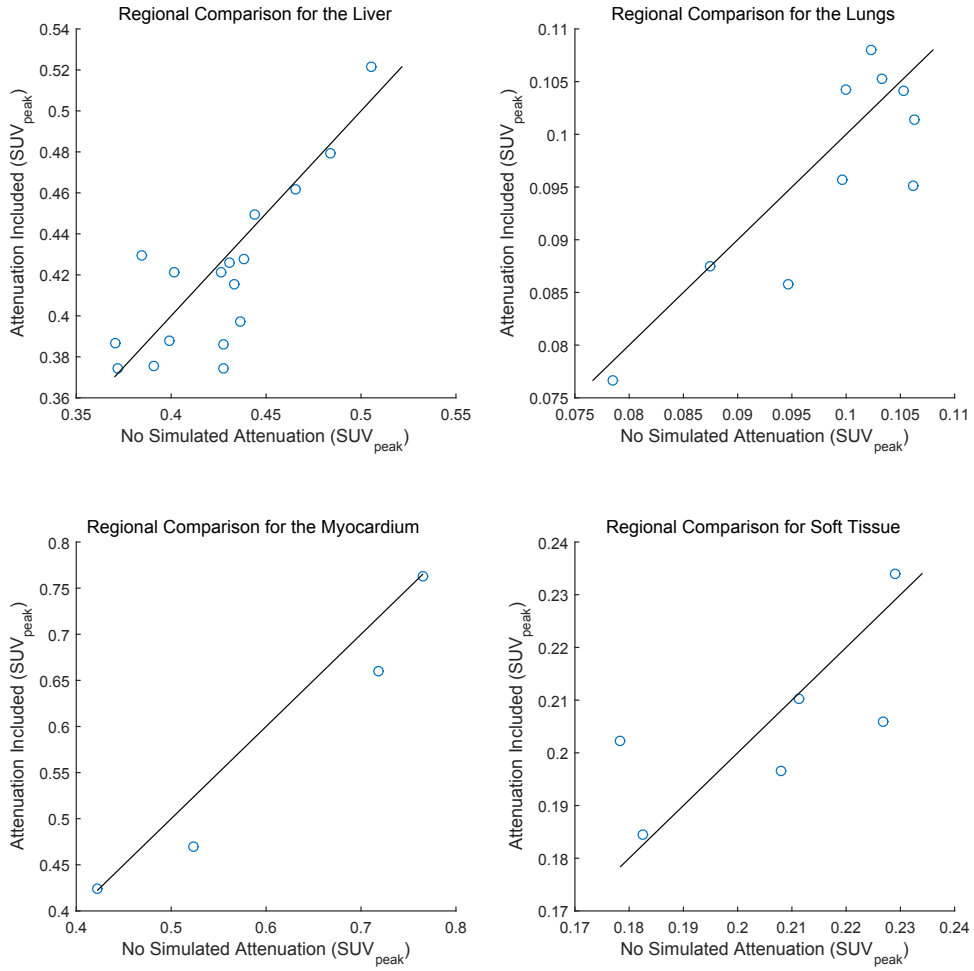

Figure 8.7: **Scatter plots of individual lesion measurements for regional analysis** The individual lesion $\mathrm{SUV}_{\mathrm{peak}}$ measurements, separated according to tissue type, are displayed in these four scatter plots. The diagonal line represents no relative change.

future with incorporation of further data corrections will also be briefly discussed, followed by a discussion of the ME-MCIR approach overall.

8.5.1 Performance with Attenuation Effects

In general, it appears that the motion estimation technique was still able to motion correct lesions after including attenuation effects into the simulation and the estimation procedure. A small drop in the lesion peak uptake of $-1.6 \pm 5.9\%$ was observed when comparing the attenuation-included data to the emission-only data, although it is unclear if this was a real effect because this difference was not statistically significant, either globally ($p = 0.216$) or in specific tissues (table 8.1). Similarly, there was no statistically-significant difference between the lesion width measurements. Position recovery, however, was statistically significant, suggesting that there was a small average displacement of 2.0 ± 2.9 mm in the centre of attenuation-included lesions compared to emission-only data.

Some more qualitative results suggest that there could be an underlying problem, however. Some attenuation artefacts were noticed in some images, usually above the left hemidiaphragm. To add to this, there were clearly some disparities between the respiratory signal estimates that suggest some lack of reliability in the attenuation-only case. For example, the estimates of s in the emission-only reconstructions in chapter 7 were approximately evenly-spread between the minimum and maximum values observed. However, in the attenuation-included reconstructions, the estimates of s were not evenly spaced.

There are several possible explanations that could explain this difference, which will require further research. Firstly, there are limitations in the simulations: only 18 motion fields were available to simulate intra-gate motion effects. With 6 gates, this averages at 3 per gate, although this varies depending on the distribution of breathing amplitudes in the dataset. Therefore, one possibility is that the distribution of intra-gate motion estimates were not smooth enough in some gates. This could be addressed by creating any number of images in intermediate breathing positions by interpolating between those observed. However, for more realistic breathing data,

it would be best to use a large number of dynamic MR images from each volunteer to ensure a reliable motion blur is present in each gate, or perhaps even trying the method on real PET-MR data.

Alternatively, there could be some ill-posedness of the estimation problem: high-contrast edges generated by the attenuation map might artificially appear to be the same higher-contrast edges expected (e.g. for the diaphragm) in accurate motion correction. Whilst it is not immediately clear from the information available whether this is the case, it would be interesting to try other forms of μ-map, such as one filtered to match the PET scanner resolution, as is often done in clinical PET. The simulations could also be used to vary relative contrast of different tissues (both in the attenuation map and emission map) to see whether any differences can be observed. If the high-contrast edges do prove to be the problem, a more robust optimisation approach could be employed to avoid local maxima, such as preconditioning or incorporation into an EM-style algorithm.

In section 7.2.2, some simplifying assumptions were made when deriving the motion estimation algorithm. At the time, it was assumed that the effects of warped basis functions and mass preservation might be unimportant in a practical setting. However, failing to account for mass preservation could lead to discrepancies in apparent activity distribution. This would have been emphasised when including attenuation since it involves an exponent: small inconsistencies in attenuation coefficient can become significant. An improved estimation gradient should take this mass preservation into account, although since the mass preservation term would also depend on the motion parameters being estimated, incorporation might not be straightforward.

Finally, although some initial testing was done to decide on the best VOI for the motion estimation, the estimation was found to be quite sensitive to the size and positioning of the window. A window placed over the right lung and liver was found to yield the most reliable results, possibly because this is the region where the correlation between the signal and the overall motion is strongest. However, there might need to be some fine-tuning to optimise this for better reliability. Another practical solution could be the introduction of some regularisation into the motion estimation by enforcing some constraints on the relative spacing of the respiratory

signal estimates. The downside of this is the addition of free parameters when an aim of this approach is to minimise on this.

Despite the above, solving this problem might not be of major importance. Certainly, from a practical viewpoint, the lack of statistically-significant differences from the emission-only version of ME-MCIR suggests that the inclusion of attenuation still performs as well in terms of lesion recovery and characterisation. This still has all of the features that make the method useful, such as robustness to noise, although more data and testing would be beneficial to understand whether a statistically-significant difference can be found, and if so, why.

8.5.2 Discussion of the ME-MCIR Algorithm Overall and Future Improvements

To conclude this chapter, the joint estimation method introduced in this thesis will be discussed in generality. First, the improvement of the data model to handle further effects will be discussed. The advantages and disadvantages of the ME-MCIR approach will be discussed. Challenges to the approach will be outlined, with some suggestions on ways in which it could be improved.

Including Further Effects into the Data Model

Although this chapter has sought to improve the model from an emission-only data model to one that includes attenuation, there are still major effects that might need to be accounted for before reconstruction can be used on real patient data. On the other hand, it is unknown how much these will negatively affect estimation of the motion parameters.

For example, attenuation and its correction involves estimation of a μ-map. This image is highly dependent on motion itself, but also strongly affects the appearance of the image estimate, and as such is an important quality that can affect motion estimation. Incorporation was successful in terms of lesion quantification, although qualitative evidence suggests that local minima might exist.

On the bright side, this might not be the case for scatter and random estimation, since both are relatively low-resolution effects and might prove to be insensitive to slight variations in the subject's position over the course of PET data acquisition. If this is the case, it might suffice to use 'static' random and scatter estimates during reconstruction, with only some additional consideration to see how they must be included in the update equation for s.

Timing and Convergence Speed

The algorithm is relatively slow since it requires many forward projections: in addition to each MCIR-type image update, each AF sinogram must be recalculated from the attenuation maps (requiring an additional forward-projection per gate) and 3 forward-projections per internal variable must take place as part of the motion update. As the algorithm stands with the inclusion with attenuation, 60% of each iteration is taken up by calls to the forward-projector, and 13% by the back-projector. Reliance on this number of projections could be reduced by, for example, only updating the motion operator every 3 iterations of image updates. Alternatively, or in addition, the motion estimation could halted once the update gradient falls within an interval of accepted error. This is supported by the fact that the motion parameters appear to converge relatively quickly (within approximately 20 iterations).

The gradient-ascent stepsize β was selected to set a sensible scale of the problem (discussed in section 7.3.4). This could be optimised more formally using a controlled study, or by employing more sophisticated methods such as conjugate gradient methods. Ultimately it would be tempting to develop the motion estimation into an EM-type optimisation through a complete-data formulation. Finally, the interleaved activity/motion estimation is quite ad-hoc. It would be interesting to incorporate this into a theoretically-rigorous framework, such as a MAP framework.

Motion Modelling

The motion model used to test this algorithm was relatively simplistic. It only required 1 motion-correlated input parameter to produce motion estimates for a PET gate. This level of restriction is a double-edged sword. It is useful because the entire image can be visualised as being 'on rails'; that is, one number controls the position of every voxel at once. This parameter reduction allows significant noise-averaging, giving it robustness to high-noise motion estimation.

Conversely, the relationship between the input parameter and the global motion estimate varied in reliability. In this case, the use of a diaphragmatic navigator meant good local correspondence around the liver. Correlation with motion of the organs in the left thorax was not as strong and any variation in this location was not accounted for, possibly explaining the attenuation artefacts observed in this region (figure 8.4).

This could be accounted for by using a more complex motion model. For example, two motion-correlated signals could be used to produce motion estimates (such as a navigator across the diaphragm below each lung). This might cause issues since the search space for optimisation becomes much larger for each gate, potentially removing the technique's robustness to noise. Alternatively, a pseudo-1D hidden variable could be used, which can search a 2D parameter space but is constrained to remain close to a line, such as the technique used in [177].

It would be advantageous to find a better MR-based method of motion estimation for the motion model formation. The current implementation, which uses lreg to register rapidly-acquired 3D MR volumes, appears to struggle with reliable motion estimation in the lungs due to lack of information to the registration algorithm. The work in [129] might prove to be a solution to this since it can create MR images with clear pulmonary vascular structure, providing much more information in these regions.

Finally, motion modelling appears to be a powerful method for practical applications. Motion estimates derived from the MR are able to estimate types of motion that PET alone is unable to do, such as twisting and sliding [108]. The use of motion

modelling as a constraint is also attractive from a practical setting: the constrained set of outcomes will always be limited to what has been observed.

A disadvantage of this is that it still requires an amount of MR time to obtain the data for forming the model, and MR acquisition protocols in PET-MR are already short on time [183]. The current method requires relatively little MR acquisition time, at approximately 30 s. In the future it may be possible to build and employ population-based models, although these can be unreliable. This still requires much research and development, especially for application in a PET setting where irregular breathing and morphology due to disease are common.

Further disadvantages related to motion modelling and this technique involve the assumptions made to apply it. For example, it is assumed that the end-expiration position observed in the MR dynamic acquisition will match the corresponding PET gate. This was selected since it is the most repeatable position [41]. However, this may not be the case – displacements due to bulk motion would invalidate this assumption, for example. Furthermore, this would not matter if registering PET to itself directly.

Related to the above assumption is that the average cycle motion described by the model is assumed to be sufficient to capture the motion in the PET gates. It is possible, for example, that a subject's breathing pattern might change significantly after the acquiring the MR data for model formation.

The present formulation has no way of taking these mismatches into account, and could lead to nonsensical solutions to the motion estimation problem. One way of minimising these effects could be but could be addressed by using the Bayesian approach discussed above – this softens the trajectory that the motion model constrains to and allows a small degree of extrapolation. Adaptive motion models can also be used to take changes in breathing pattern into account.

8.6 Conclusion

In this chapter, the ME-MCIR algorithm was extended to include attenuation in the model of the mean data and assessed using PET simulations based on real MR-measured motion estimates and anatomy. The experimental results suggest that the correction term in the update equation successfully compensates for the introduction of attenuation into the data, although there remain some questions about how well posed the estimation of multiple motion parameters is. This is thought to be due to the appearance of features of image introduced by attenuation correction, such as high contrast edges before good motion estimates for correction have been identified. Overall the ME-MCIR algorithm remains a promising method for motion correcting PET images, although there still remain a number of improvements that are required – both for application to patient data, and in general.

Chapter 9

Conclusion

9.1 Summary of Thesis Achievements

In this thesis, methods driving parameterised motion models with PET data were assessed. This began with a proof-of-principle approach, which formed a 3D subject-specific MR motion model from a short dynamic MR acquisition. This model related a full voxelwise motion description with a strongly correlated parameter, the internal displacement of the right diaphragm. The model provided a way to constrain the registration of lower-count PET gates to biomechanically observed possibilities only. The reduced parameterisation also allowed for mitigation of the low SNR in the gates. These were then combined using RTA, which appeared to work well compared to a ground truth and a similar method driven with direct measurements of the motion parameter.

Reconstruction of low-count PET data can introduce quantification problems into the PET image, such as increased bias. To avoid this, and due to the naturalness of combining multiple estimation algorithms into one, work was undertaken to introduce this motion model into the PET image reconstruction via an MCIR-type scheme, where only one image estimate is produced by each iterative update. The motion model was incorporated by deriving an analytical gradient from the Poisson log-likelihood, which was maximised by gradient ascent. The incorporation of the motion model allowed full voxelwise motion estimation from each PET gate, but

reduced the number of parameters required for this from $J + 3JG$ to $J + G$, for J voxels and G gates. This method significantly improved quantification of FDG-avid lesions than the RTA method. The reduced parameterisation also lent robustness to noise, which appeared to improve the quantification capibility with increasing numbers of gates (and therefore fewer counts and less intra-gate motion).

A significant limitation of the approaches above is that they were developed using emission-only simulations of PET data. In an attempt to address this, and motivated by other practical factors, attenuation correction was incorporated into the objective function and its motion gradient. Incorporation of attenuation into both the simulation and motion estimation produced images whose quantification was not statistically different compared to the emission-only data and estimation method. However, there appear to be some issues in the motion estimation, possibly due to the creation of local maxima by the image qualities associated with both attenuation and motion, such as higher-contrast edges, which gradient ascent was unable to avoid.

9.2 Future Work

The joint estimation technique presented in this thesis has a number of directions it can be taken in for interesting research. For example, more work to incorporate other sources of error is likely required before it can be reliably applied to real data. Similarly, the algorithm requires a significant number of additional projections to perform the motion update for each gate. There are ways this could be decreased, either by early stopping of the motion estimate once it falls to, e.g., half of the PET resolution or by using a faster optimisation method in place of gradient ascent, such as conjugate gradient ascent. This may also lend robustness to the local minima generated by the addition of attenuation into the data model. Ideally, the need for stepsize fixing in the gradient would be removed, and as such it would be interesting to see whether this technique can be incorporated into an EM update. Finally, there is scope for increased complexity in the motion modelling. This could involve approaches that take into account variation, or by moving towards a population-

based model to remove the requirement for MR acquisition entirely. The wide range of ways to model the motion is possibly the most open area for further research.

Bibliography

[1] A. King, C. Tsoumpas, C. Buerger, V. Schulz, P. Marsden, and T. Schaeffte. Real-time respiratory motion correction for simultaneous PET-MR using an MR-derived motion model. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE*, pages 3589–3594. IEEE, 2011.

[2] K. Thielemans, C. Tsoumpas, S. Mustafovic, T. Beisel, P. Aguiar, N. Dikaios, and M. W. Jacobson. STIR: software for tomographic image reconstruction release 2. *Physics in Medicine and Biology*, 57(4):867, 2012.

[3] D. R. Balfour, P. K. Marsden, I. Polycarpou, C. Kolbitsch, and A. P. King. Respiratory motion correction of PET using MR-constrained PET-PET registration. *Biomedical Engineering Online*, 14(1):1, 2015.

[4] J. R. McClelland, D. J. Hawkes, T. Schaeffter, and A. P. King. Respiratory motion models: a review. *Medical Image Analysis*, 17(1):19–42, 2013.

[5] T. Li, B. Thorndyke, E. Schreibmann, Y. Yang, and L. Xing. Model-based image reconstruction for four-dimensional PET. *Medical Physics*, 33(5):1288–1298, 2006.

[6] S. R. Cherry, J. A. Sorenson, and M. E. Phelps. *Physics in nuclear medicine*. Elsevier Health Sciences, 2012.

[7] R. Chandra and D. Mihailidis. *Nuclear medicine physics: the basics*. Wolters Kluwer Health/Lippincott Williams & Wilkins, 2012.

[8] G. Delso and S. Ziegler. PET/MRI system design. *European Journal of Nuclear Medicine and Molecular Imaging*, 36(1):86–92, 2009.

[9] D. L. Bailey. Data acquisition and performance characterization in PET. In *Positron emission tomography*, pages 41–62. Springer, 2005.

[10] H. Zaidi, M.-L. Montandon, and D. O. Slosman. Magnetic resonance imaging-guided attenuation and scatter corrections in three-dimensional brain positron emission tomography. *Medical physics*, 30(5):937–948, 2003.

[11] A. Martinez-Möller, M. Souvatzoglou, G. Delso, R. A. Bundschuh, C. Chefd'hotel, S. I. Ziegler, N. Navab, M. Schwaiger, and S. G. Nekolla. Tissue classification as a potential approach for attenuation correction in whole-body PET/MRI: evaluation with PET/CT data. *Journal of Nuclear Medicine*, 50(4):520–526, 2009.

[12] V. Keereman, Y. Fierens, T. Broux, Y. De Deene, M. Lonneux, and S. Vandenberghe. MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences. *Journal of Nuclear Medicine*, 51(5):812–818, 2010.

[13] N. Burgos, M. J. Cardoso, M. Modat, S. Pedemonte, J. Dickson, A. Barnes, J. S. Duncan, D. Atkinson, S. R. Arridge, B. F. Hutton, et al. Attenuation correction synthesis for hybrid PET-MR scanners. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*, pages 147–154. Springer, 2013.

[14] S. R. Cherry and M. Dahlbom. Pet: physics, instrumentation, and scanners. In *PET*, pages 1–117. Springer, 2006.

[15] C. C. Watson. New, faster, image-based scatter correction for 3D PET. *Nuclear Science, IEEE Transactions on*, 47(4):1587–1594, 2000.

[16] J. Qi and R. M. Leahy. Iterative reconstruction techniques in emission computed tomography. *Physics in Medicine and Biology*, 51(15):R541, 2006.

[17] A. Rahmim, J. Qi, and V. Sossi. Resolution modeling in PET imaging: theory, practice, benefits, and pitfalls. *Medical Physics*, 40(6):064301, 2013.

[18] M. Soret, S. L. Bacharach, and I. Buvat. Partial-volume effect in PET tumor imaging. *Journal of Nuclear Medicine*, 48(6):932–945, 2007.

[19] D. Daou. Respiratory motion handling is mandatory to accomplish the high-resolution PET destiny. *European Journal of Nuclear Medicine and Molecular Imaging*, 35(11):1961–1970, 2008.

[20] I. Polycarpou, C. Tsoumpas, A. P. King, and P. K. Marsden. Impact of respiratory motion correction and spatial resolution on lesion detection in PET: a simulation study based on real mr dynamic data. *Physics in Medicine and Biology*, 59(3):697, 2014.

[21] P. E. Kinahan and J. Rogers. Analytic 3D image reconstruction using all detected events. *Nuclear Science, IEEE Transactions on*, 36(1):964–968, 1989.

[22] M. Krzywinski, V. Sossi, and T. Ruth. Comparison of FORE, OSEM and SAGE algorithms to 3DRP in 3D PET using phantom and human subject data. *Nuclear Science, IEEE Transactions on*, 46(4):1114–1120, 1999.

[23] M. Defrise. A factorization method for the 3D X-ray transform. *Inverse Problems*, 11(5):983, 1995.

[24] R. M. Lewitt. Alternatives to voxels for image representation in iterative reconstruction algorithms. *Physics in Medicine and Biology*, 37(3):705, 1992.

[25] A. Entezari, M. Nilchian, and M. Unser. A box spline calculus for the discretization of computed tomography reconstruction problems. *Medical Imaging, IEEE Transactions on*, 31(8):1532–1541, 2012.

[26] S. Matej and R. M. Lewitt. Practical considerations for 3-D image reconstruction using spherically symmetric volume elements. *Medical Imaging, IEEE Transactions on*, 15(1):68–78, 1996.

[27] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B (Methodological)*, pages 1–38, 1977.

[28] L. A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *Medical Imaging, IEEE Transactions on*, 1(2):113–122, 1982.

[29] H. M. Hudson and R. S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *Medical Imaging, IEEE Transactions on*, 13(4):601–609, 1994.

[30] J. A. Fessler and A. O. Hero. Space-alternating generalized expectation-maximization algorithm. *Signal Processing, IEEE Transactions on*, 42(10):2664–2677, 1994.

[31] J. Browne and A. R. De Pierro. A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography. *Medical Imaging, IEEE Transactions on*, 15(5):687–699, 1996.

[32] S. Geman and D. McClure. Bayesian image analysis: an application to single photon emission tomography. *American Statististical Association*, pages 12–18, 1985.

[33] P. J. Green. Bayesian reconstructions from emission tomography data using a modified EM algorithm. *Medical Imaging, IEEE Transactions on*, 9(1):84–93, 1990.

[34] C. Kolbitsch, C. Prieto, C. Tsoumpas, and T. Schaeffter. A 3D MR-acquisition scheme for nonrigid bulk motion correction in simultaneous PET-MR. *Medical Physics*, 41(8):082304, 2014.

[35] R. Fulton, L. Tellmann, U. Pietrzyk, O. Winz, I. Stangier, I. Nickel, A. Schmid, S. Meikle, and H. Herzog. Accuracy of motion correction methods for PET brain imaging. In *Nuclear Science Symposium Conference Record, 2004 IEEE*, volume 7, pages 4226–4230. IEEE, 2004.

[36] Y. E. Erdi, S. A. Nehmeh, T. Pan, A. Pevsner, K. E. Rosenzweig, G. Mageras, E. D. Yorke, H. Schoder, W. Hsiao, O. D. Squire, et al. The CT motion quantitation of lung lesions and its impact on PET-measured SUVs. *Journal of Nuclear Medicine*, 45(8):1287–1292, 2004.

[37] Z. Cao, D. R. Gilland, B. A. Mair, and R. J. Jaszczak. 3D motion estimation with image reconstruction for gated cardiac ECT. In *Nuclear Science Sym-*

posium Conference Record, 2002 IEEE, volume 3, pages 1869–1873. IEEE, 2002.

[38] Y. Nakamoto, B. B. Chin, C. Cohade, M. Osman, M. Tatsumi, and R. L. Wahl. PET/CT: artifacts caused by bowel motion. *Nuclear Medicine Communications*, 25(3):221–225, 2004.

[39] K. Konno and J. Mead. Measurement of the separate volume changes of rib cage and abdomen during breathing. *Journal of Applied Physiology*, 22(3):407–422, 1967.

[40] G. Benchetrit. Breathing pattern in humans: diversity and individuality. *Respiration Physiology*, 122(2):123–129, 2000.

[41] J. Blackall, S. Ahmad, M. Miquel, J. McClelland, D. Landau, and D. Hawkes. MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning. *Physics in Medicine and Biology*, 51(17):4147, 2006.

[42] F. Ernst, R. Bruder, A. Schlaefer, and A. Schweikard. Correlation between external and internal respiratory motion: a validation study. *International Journal of Computer Assisted Radiology and Surgery*, 7(3):483–492, 2012.

[43] J. Escolar and A. Escolar. Lung hysteresis: a morphological view. *Histology and Histopathology*, 19(1):159–166, 2004.

[44] A. De Troyer and M. Estenne. Coordination between rib cage muscles and diaphragm during quiet breathing in humans. *Journal of Applied Physiology*, 57(3):899–906, 1984.

[45] M. M. Osman, C. Cohade, Y. Nakamoto, L. T. Marshall, J. P. Leal, and R. L. Wahl. Clinically significant inaccurate localization of lesions with PET/CT: frequency in 300 patients. *Journal of Nuclear Medicine*, 44(2):240–243, 2003.

[46] C. W. Stevens, R. F. Munden, K. M. Forster, J. F. Kelly, Z. Liao, G. Starkschall, S. Tucker, and R. Komaki. Respiratory-driven lung tumor motion is independent of tumor size, tumor location, and pulmonary function.

International Journal of Radiation Oncology Biology* Physics*, 51(1):62–68, 2001.

[47] Y. Seppenwoolde, H. Shirato, K. Kitamura, S. Shimizu, M. van Herk, J. V. Lebesque, and K. Miyasaka. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *International Journal of Radiation Oncology* Biology* Physics*, 53(4):822–834, 2002.

[48] C. Liu, L. A. Pierce II, A. M. Alessio, and P. E. Kinahan. The impact of respiratory motion on tumor quantification and delineation in static PET/CT imaging. *Physics in Medicine and Biology*, 54(24):7345, 2009.

[49] T. Kawano, E. Ohtake, and T. Inoue. Deep-inspiration breath-hold PET/CT versus free breathing PET/CT and respiratory gating PET for reference: evaluation in 95 patients with lung cancer. *Annals of Nuclear Medicine*, 25(2):109–116, 2011.

[50] D. Papathanassiou, S. Becker, R. Amir, B. Menéroux, and J.-C. Liehn. Respiratory motion artefact in the liver dome on FDG PET/CT: comparison of attenuation correction with CT and a caesium external source. *European Journal of Nuclear Medicine and Molecular Imaging*, 32(12):1422–1428, 2005.

[51] S. A. Nehmeh and Y. E. Erdi. Respiratory motion in positron emission tomography/computed tomography: a review. In *Seminars in Nuclear Medicine*, volume 38, pages 167–176. Elsevier, 2008.

[52] R. R. Fulton, S. R. Meikle, S. Eberl, J. Pfeiffer, and C. J. Constable. Correction for head movements in positron emission tomography using an optical motion-tracking system. *Nuclear Science, IEEE Transactions on*, 49(1):116–123, 2002.

[53] P. J. Noonan, J. Howard, T. F. Cootes, W. A. Hallett, and R. Hinz. Real-time markerless rigid body head motion tracking using the microsoft kinect. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE*, pages 2241–2246. IEEE, 2012.

[54] C. Ozturk, J. A. Derbyshire, and E. McVeigh. Estimating motion from MRI data. *Proceedings of the IEEE*, 91(10):1627–1648, 2003.

[55] C. Würslin, H. Schmidt, P. Martirosian, C. Brendle, A. Boss, N. F. Schwenzer, and L. Stegger. Respiratory motion correction in oncologic PET using T1-weighted MR imaging on a simultaneous whole-body PET/MR system. *Journal of Nuclear Medicine*, 54(3):464–471, 2013.

[56] M. Zaitsev, J. Maclaren, and M. Herbst. Motion artifacts in MRI: A complex problem with many partial solutions. *Journal of Magnetic Resonance Imaging*, 42(4):887–901, 2015.

[57] R. Mezrich. A perspective on k-space. *Radiology*, 195(2):297–315, 1995. PMID: 7724743.

[58] M. L. Wood and R. M. Henkelman. MR image artifacts from periodic motion. *Medical Physics*, 12(2):143–151, 1985.

[59] K. Nehrke and P. Börnert. Prospective correction of affine motion for arbitrary MR sequences on a clinical scanner. *Magnetic Resonance in Medicine*, 54(5):1130–1138, 2005.

[60] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger. SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine*, 42(5):952–962, 1999.

[61] M. Usman, D. Atkinson, F. Odille, C. Kolbitsch, G. Vaillant, T. Schaeffter, P. G. Batchelor, and C. Prieto. Motion corrected compressed sensing for free-breathing dynamic cardiac MRI. *Magnetic Resonance in Medicine*, 70(2):504–516, 2013.

[62] R. Grimm, S. Fürst, M. Souvatzoglou, C. Forman, J. Hutter, I. Dregely, S. I. Ziegler, B. Kiefer, J. Hornegger, K. T. Block, et al. Self-gated MRI motion modeling for respiratory motion compensation in integrated PET/MRI. *Medical Image Analysis*, 19(1):110–120, 2015.

[63] P. Batchelor, D. Atkinson, P. Irarrazaval, D. Hill, J. Hajnal, and D. Larkman. Matrix description of general motion correction applied to multishot images. *Magnetic Resonance in Medicine*, 54(5):1273–1280, 2005.

[64] C. R. Crawford, K. F. King, C. J. Ritchie, and J. D. Godwin. Respiratory compensation in projection imaging using a magnification and displacement model. *IEEE Transactions on Medical Imaging*, 15(3):327–332, 1996.

[65] D. Visvikis, D. C. Costa, I. Croasdale, A. H. R. Lonn, J. Bomanji, S. Gacinovic, and P. J. Ell. CT-based attenuation correction in the calculation of semi-quantitative indices of [18F]FDG uptake in PET. *European Journal of Nuclear Medicine and Molecular Imaging*, 30(3):344–353, 2003.

[66] C. Cohade, M. Osman, L. T. Marshall, and R. L. Wahl. PET-CT: accuracy of PET and CT spatial registration of lung lesions. *European Journal of Nuclear Medicine and Molecular Imaging*, 30(5):721–726, 2003.

[67] S. Vedam, P. Keall, V. Kini, H. Mostafavi, H. Shukla, and R. Mohan. Acquiring a four-dimensional computed tomography dataset using an external respiratory signal. *Physics in Medicine and Biology*, 48(1):45, 2002.

[68] C. Buerger, C. Tsoumpas, A. Aitken, A. P. King, P. Schleyer, V. Schulz, P. K. Marsden, and T. Schaeffter. Investigation of MR-based attenuation correction and motion compensation for hybrid PET/MR. *Nuclear Science, IEEE Transactions on*, 59(5):1967–1976, 2012.

[69] Y. Picard and C. J. Thompson. Motion correction of PET images using multiple acquisition frames. *Medical Imaging, IEEE Transactions on*, 16(2):137–144, 1997.

[70] S. A. Nehmeh, Y. E. Erdi, G. S. Meirelles, O. Squire, S. M. Larson, J. L. Humm, and H. Schöder. Deep-inspiration breath-hold PET/CT of the thorax. *Journal of Nuclear Medicine*, 48(1):22–26, 2007.

[71] M. Dawood, F. Büther, L. Stegger, X. Jiang, O. Schober, M. Schäfers, and K. P. Schäfers. Optimal number of respiratory gates in positron emission tomography: a cardiac patient study. *Medical physics*, 36(5):1775–1784, 2009.

[72] F. Odille, P.-A. Vuissoz, P.-Y. Marie, and J. Felblinger. Generalized reconstruction by inversion of coupled systems (GRICS) applied to free-breathing MRI. *Magnetic Resonance in Medicine*, 60(1):146–157, 2008.

[73] J. D. Hoisak, K. E. Sixel, R. Tirona, P. C. Cheung, and J.-P. Pignol. Correlation of lung tumor motion with external surrogate indicators of respiration. *International Journal of Radiation Oncology* Biology* Physics*, 60(4):1298–1306, 2004.

[74] L. Boucher, S. Rodrigue, R. Lecomte, and F. Bénard. Respiratory gating for 3-dimensional PET of the thorax: feasibility and initial results. *Journal of Nuclear Medicine*, 45(2):214–219, 2004.

[75] M. Dawood, F. Büther, X. Jiang, and K. P. Schäfers. Respiratory motion correction in 3-D PET data with advanced optical flow algorithms. *Medical Imaging, IEEE Transactions on*, 27(8):1164–1175, 2008.

[76] R. Zeng, J. A. Fessler, J. M. Balter, and P. A. Balter. Iterative sorting for four-dimensional CT images based on internal anatomy motion. *Medical physics*, 35(3):917–926, 2008.

[77] T. A. Hope, E. F. Verdin, E. K. Bergsland, M. A. Ohliger, C. U. Corvera, and E. K. Nakakura. Correcting for respiratory motion in liver PET/MRI: preliminary evaluation of the utility of bellows and navigated hepatobiliary phase imaging. *EJNMMI physics*, 2(1):1–11, 2015.

[78] P. G. Danias, M. V. McConnell, V. C. Khasgiwala, M. L. Chuang, R. R. Edelman, and W. J. Manning. Prospective navigator correction of image position for coronary MR angiography. *Radiology*, 203(3):733–736, 1997.

[79] B. S. Kim, J. H. Kim, G. M. Choi, S. H. Kim, J. K. Park, B.-C. Song, and W. Kang. Comparison of three free-breathing T2-weighted MRI sequences in the evaluation of focal liver lesions. *American Journal of Roentgenology*, 190(1):W19–W27, 2008.

[80] G. J. Klein, B. W. Reutter, and R. H. Huesman. Four-dimensional affine registration models for respiratory-gated PET. *Nuclear Science, IEEE Transactions on*, 48(3):756–760, 2001.

[81] R. A. Bundschuh, A. Martínez-Möller, M. Essler, S. G. Nekolla, S. I. Ziegler, and M. Schwaiger. Local motion correction for lung tumours in PET/CT: first results. *European Journal of Nuclear Medicine and Molecular Imaging*, 35(11):1981–1988, 2008.

[82] D. Visvikis, O. Barret, T. Fryer, A. Turzo, F. Lamare, L. Rest, C. Cheze, and Y. Bizais. A posteriori respiratory motion gating of dynamic PET images. In *Nuclear Science Symposium Conference Record, 2003 IEEE*, volume 5, pages 3276–3280. IEEE, 2003.

[83] P. J. Schleyer, M. J. O'Doherty, S. F. Barrington, and P. K. Marsden. Retrospective data-driven respiratory gating for PET/CT. *Physics in Medicine and Biology*, 54(7):1935, 2009.

[84] F. Büther, M. Dawood, L. Stegger, F. Wübbeling, M. Schäfers, O. Schober, and K. P. Schäfers. List mode–driven cardiac and respiratory gating in PET. *Journal of Nuclear Medicine*, 50(5):674–681, 2009.

[85] K. Thielemans, S. Rathore, F. Engbrant, and P. Razifar. Device-less gating for PET/CT using PCA. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE*, pages 3904–3910. IEEE, 2011.

[86] K. Thielemans, P. Schleyer, P. K. Marsden, R. M. Manjeshwar, S. D. Wollenweber, and A. Ganin. Comparison of different methods for data-driven respiratory gating of PET data. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2013 IEEE*, pages 1–4. IEEE, 2013.

[87] D. Visvikis, F. Lamare, P. Bruyant, N. Boussion, and C. C. L. Rest. Respiratory motion in positron emission tomography for oncology applications: Problems and solutions. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 569(2):453 – 457, 2006. Proceedings of the 3rd International Conference on

Imaging Technologies in Biomedical SciencesInnovation in Nuclear and Radiological Imaging: from Basic Research to Clinical Application.

[88] R. Manber, K. Thielemans, B. F. Hutton, S. Wan, J. McClelland, A. Barnes, S. Arridge, S. Ourselin, and D. Atkinson. Joint PET-MR respiratory motion models for clinical PET motion correction. *Physics in Medicine and Biology*, 61(17):6515, 2016.

[89] C. Liu, A. Alessio, L. Pierce, K. Thielemans, S. Wollenweber, A. Ganin, and P. Kinahan. Quiescent period respiratory gating for pet/ct. *Medical Physics*, 37(9):5037–5043, 2010.

[90] W. Grootjans, F. Tixier, C. S. van der Vos, D. Vriens, C. C. Le Rest, J. Bussink, W. J. Oyen, L.-F. de Geus-Oei, D. Visvikis, and E. P. Visser. The impact of optimal respiratory gating and image noise on evaluation of intratumor heterogeneity on 18F-FDG PET imaging of lung cancer. *Journal of Nuclear Medicine*, 57(11):1692–1698, 2016.

[91] E. Rietzel, G. T. Chen, N. C. Choi, and C. G. Willet. Four-dimensional image-based treatment planning: Target volume segmentation and dose calculation in the presence of respiratory motion. *International Journal of Radiation Oncology* Biology* Physics*, 61(5):1535–1550, 2005.

[92] O. Camara, G. Delso, and I. Bloch. Evaluation of a thoracic elastic registration method using anatomical constraints in oncology. In *2nd. Joint Conference of the IEEE Engineering in Medicine and Biology Society*. Citeseer, 2002.

[93] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank. PET-CT image registration in the chest using free-form deformations. *Medical Imaging, IEEE Transactions on*, 22(1):120–128, 2003.

[94] F. Lamare, M. L. Carbayo, T. Cresson, G. Kontaxakis, A. Santos, C. C. Le Rest, A. Reader, and D. Visvikis. List-mode-based reconstruction for respiratory motion correction in PET using non-rigid body transformations. *Physics in Medicine and Biology*, 52(17):5187, 2007.

[95] W. Bai and M. Brady. Regularized B-spline deformable registration for respiratory motion correction in PET images. *Physics in Medicine and Biology*, 54(9):2719, 2009.

[96] S. Y. Chun, T. G. Reese, J. Ouyang, B. Guerin, C. Catana, X. Zhu, N. M. Alpert, and G. El Fakhri. MRI-based nonrigid motion correction in simultaneous PET/MRI. *Journal of Nuclear Medicine*, 53(8):1284–1291, 2012.

[97] A. Bousse, O. Bertolli, D. Atkinson, S. Arridge, S. Ourselin, B. Hutton, and K. Thielemans. Maximum-likelihood joint image reconstruction/motion estimation in attenuation-corrected respiratory gated PET/CT using a single attenuation map. *IEEE Transactions on Medical Imaging*, 2016.

[98] J. A. Schnabel, D. Rueckert, M. Quist, J. M. Blackall, A. D. Castellano-Smith, T. Hartkens, G. P. Penney, W. A. Hall, H. Liu, C. L. Truwit, F. A. Gerritsen, D. L. G. Hill, and D. J. Hawkes. *A Generic Framework for Non-rigid Registration Based on Non-uniform Multi-level Free-Form Deformations*, volume 2208, pages 573–581. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[99] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache. Symmetric log-domain diffeomorphic registration: A demons-based approach. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008*, pages 754–761. Springer, 2008.

[100] W. R. Crum, O. Camara, and D. J. Hawkes. Methods for inverting dense displacement fields: evaluation in brain image registration. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 900–907. Springer, 2007.

[101] K. Thielemans, S. Mustafovic, and L. Schnorr. Image reconstruction of motion corrected sinograms. In *Nuclear Science Symposium Conference Record, 2003 IEEE*, volume 4, pages 2401–2406. IEEE, 2003.

[102] P. M. Bloomfield, T. J. Spinks, J. Reed, L. Schnorr, A. M. Westrip, L. Livieratos, R. Fulton, and T. Jones. The design and implementation of a motion

correction scheme for neurological PET. *Physics in Medicine and Biology*, 48(8):959, 2003.

[103] L. Livieratos, L. Stegger, P. Bloomfield, K. Schafers, D. Bailey, and P. Camici. Rigid-body transformation of list-mode projection data for respiratory motion correction in cardiac PET. *Physics in Medicine and Biology*, 50(14):3313, 2005.

[104] F. Lamare, T. Cresson, J. Savean, C. C. Le Rest, A. Reader, and D. Visvikis. Respiratory motion correction for PET oncology applications using affine transformation of list mode data. *Physics in Medicine and Biology*, 52(1):121, 2006.

[105] C. Catana, T. Benner, A. van der Kouwe, L. Byars, M. Hamm, D. B. Chonde, C. J. Michel, G. El Fakhri, M. Schmand, and A. G. Sorensen. MRI-assisted PET motion correction for neurologic studies in an integrated MR-PET scanner. *Journal of Nuclear Medicine*, 52(1):154–161, 2011.

[106] S. Pedemonte, A. Bousse, B. F. Hutton, S. Arridge, and S. Ourselin. 4-D generative model for PET/MRI reconstruction. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*, pages 581–588. Springer, 2011.

[107] M. Jacobson and J. A. Fessler. Joint estimation of image and deformation parameters in motion-corrected PET. In *Nuclear Science Symposium Conference Record, 2003 IEEE*, volume 5, pages 3290–3294. IEEE, 2003.

[108] G. Klein, B. Reutter, and R. Huesman. Non-rigid summing of gated PET via optical flow. *Nuclear Science, IEEE Transactions on*, 44(4):1509–1512, 1997.

[109] F. Gigengack, L. Ruthotto, M. Burger, C. H. Wolters, X. Jiang, and K. P. Schäfers. Motion correction in dual gated cardiac PET using mass-preserving image registration. *Medical Imaging, IEEE Transactions on*, 31(3):698–712, 2012.

[110] H. Fayad, H. Schmidt, C. Wuerslin, and D. Visvikis. Generation of 4D MR images based on 4D PET derived motion fields. In *PSMR 2013: PET-MR*

and SPECT-MR: Current status of instrumentation, applications and developments, 2013.

[111] D. R. Gilland, B. A. Mair, J. E. Bowsher, and R. J. Jaszczak. Simultaneous reconstruction and motion estimation for gated cardiac ECT. *Nuclear Science, IEEE Transactions on*, 49(5):2344–2349, 2002.

[112] B. A. Mair, D. R. Gilland, and J. Sun. Estimation of images and nonrigid deformations in gated emission CT. *Medical Imaging, IEEE Transactions on*, 25(9):1130–1144, 2006.

[113] R. Manjeshwar, X. Tao, E. Asma, and K. Thielemans. Motion compensated image reconstruction of respiratory gated PET/CT. In *Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on*, pages 674–677. IEEE, 2006.

[114] E. J. Gravier and Y. Yang. Motion-compensated reconstruction of tomographic image sequences. *Nuclear Science, IEEE Transactions on*, 52(1):51–56, 2005.

[115] E. Gravier, Y. Yang, M. A. King, and M. Jin. Fully 4D motion-compensated reconstruction of cardiac SPECT images. *Physics in Medicine and Biology*, 51(18):4603, 2006.

[116] M. Blume, A. Martinez-Möller, A. Keil, N. Navab, and M. Rafeca. Joint reconstruction of image and motion in gated positron emission tomography. *Medical Imaging, IEEE Transactions on*, 29(11):1892–1906, 2010.

[117] M. Dawood, N. Lang, X. Jiang, and K. P. Schafers. Lung motion correction on respiratory gated 3-D PET/CT images. *IEEE Transactions on Medical Imaging*, 25(4):476–485, 2006.

[118] D. R. Gilland, B. A. Mair, and J. G. Parker. Motion estimation for cardiac emission tomography by optical flow methods. *Physics in Medicine and Biology*, 53(11):2991, 2008.

[119] A. Sotiras, C. Davatzikos, and N. Paragios. Deformable medical image registration: A survey. *Medical Imaging, IEEE Transactions on*, 32(7):1153–1190, 2013.

[120] J. Vandemeulebroucke, S. Rit, J. Kybic, P. Clarysse, and D. Sarrut. Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs. *Medical physics*, 38(1):166–178, 2011.

[121] M. Blume, A. Keil, N. Navab, and M. Rafecas. Joint reconstruction of image and motion for PET: Displacement fields versus a B-spline motion model. In *Nuclear Science Symposium Conference Record (NSS/MIC), 2010 IEEE*, pages 3506–3508. IEEE, 2010.

[122] K. Thielemans, E. Asma, and R. M. Manjeshwar. Mass-preserving image registration using free-form deformation fields. In *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE*, pages 2490–2495. IEEE, 2009.

[123] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003.

[124] F. Qiao, T. Pan, J. W. Clark Jr, and O. R. Mawlawi. A motion-incorporated reconstruction method for gated PET studies. *Physics in Medicine and Biology*, 51(15):3769, 2006.

[125] J. Ehrhardt, R. Werner, D. Säring, T. Frenzel, W. Lu, D. Low, and H. Handels. An optical flow based method for improved reconstruction of 4D CT data sets acquired during free breathing. *Medical Physics*, 34(2):711–721, 2007.

[126] B. Guérin, S. Cho, S. Y. Chun, X. Zhu, N. Alpert, G. El Fakhri, T. Reese, and C. Catana. Nonrigid PET motion compensation in the lower abdomen using simultaneous tagged-MRI and PET imaging. *Medical Physics*, 38(6):3025–3038, 2011.

[127] H. Fayad, F. Odille, H. Schmidt, C. Würslin, T. Küstner, J. Felblinger, and D. Visvikis. The use of a generalized reconstruction by inversion of coupled systems (GRICS) approach for generic respiratory motion correction in PET/MR imaging. *Physics in Medicine and Biology*, 60(6):2529, 2015.

[128] C. Tsoumpas, C. Buerger, A. King, P. Mollet, V. Keereman, S. Vandenberghe, V. Schulz, P. Schleyer, T. Schaeffter, and P. Marsden. Fast generation of 4D PET-MR data from real dynamic MR acquisitions. *Physics in Medicine and Biology*, 56(20):6597, 2011.

[129] C. F. Baumgartner, C. Kolbitsch, D. R. Balfour, P. K. Marsden, J. R. McClelland, D. Rueckert, and A. P. King. High-resolution dynamic MR imaging of the thorax for respiratory motion correction of PET using groupwise manifold alignment. *Medical Image Analysis*, 18(7):939–952, 2014.

[130] R. Manber, K. Thielemans, B. F. Hutton, A. Barnes, S. Ourselin, S. Arridge, C. O'Meara, S. Wan, and D. Atkinson. Practical PET respiratory motion correction in clinical PET/MR. *Journal of Nuclear Medicine*, 56(6):890–896, 2015.

[131] P. J. Noonan, J. Howard, D. Tout, I. Armstrong, H. A. Williams, T. F. Cootes, W. A. Hallett, and R. Hinz. Accurate markerless respiratory tracking for gated whole body PET using the Microsoft Kinect. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE*, pages 3973–3974. IEEE, 2012.

[132] H. Fayad, T. Pan, O. Pradier, and D. Visvikis. Patient specific respiratory motion modeling using a 3D patient's external surface. *Medical Physics*, 39(6):3386–3395, 2012.

[133] K. Nehrke, P. Bornert, D. Manke, and J. C. Bock. Free-breathing cardiac MR imaging: Study of implications of respiratory motion—initial results 1. *Radiology*, 220(3):810–815, 2001.

[134] N. A. Ablitt, J. Gao, J. Keegan, L. Stegger, D. N. Firmin, and G.-Z. Yang. Predictive cardiac motion modeling and correction with partial least squares regression. *Medical Imaging, IEEE Transactions on*, 23(10):1315–1324, 2004.

[135] B. Cho, P. R. Poulsen, and P. J. Keall. Real-time tumor tracking using sequential kV imaging combined with respiratory monitoring: a general framework applicable to commonly used IGRT systems. *Physics in Medicine and Biology*, 55(12):3299, 2010.

[136] J. R. McClelland, J. M. Blackall, S. Tarte, A. C. Chandler, S. Hughes, S. Ahmad, D. B. Landau, and D. J. Hawkes. A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy. *Medical Physics*, 33(9):3348–3358, 2006.

[137] A. P. King, R. Boubertakh, K. S. Rhode, Y. Ma, P. Chinchapatnam, G. Gao, T. Tangcharoen, M. Ginks, M. Cooklin, J. S. Gill, et al. A subject-specific technique for respiratory motion correction in image-guided cardiac catheterisation procedures. *Medical Image Analysis*, 13(3):419–431, 2009.

[138] D. Peressutti, E. J. Rijkhorst, D. C. Barratt, G. P. Penney, and A. P. King. Estimating and resolving uncertainty in cardiac respiratory motion modelling. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 262–265, May 2012.

[139] A. P. King, C. Buerger, C. Tsoumpas, P. K. Marsden, and T. Schaeffter. Thoracic respiratory motion estimation from MRI using a statistical model and a 2-D image navigator. *Medical image analysis*, 16(1):252–264, 2012.

[140] M. Reyes, G. Malandain, P. M. Koulibaly, M. Gonzalez-Ballester, and J. Darcourt. Model-based respiratory motion compensation for emission tomography image reconstruction. *Physics in Medicine and Biology*, 52(12):3579, 2007.

[141] M. Fieseler, F. Gigengack, X. Jiang, and K. P. Schäfers. Motion correction of whole-body PET data with a joint PET-MRI registration functional. *Biomedical Engineering Online*, 13(2), 2014.

[142] W. Lu and T. R. Mackie. Tomographic motion detection and correction directly in sinogram space. *Physics in Medicine and Biology*, 47(8):1267, 2002.

[143] C. Tsoumpas, J. E. Mackewn, P. Halsted, A. P. King, C. Buerger, J. J. Totman, T. Schaeffter, and P. K. Marsden. Simultaneous PET–MR acquisition and MR-derived motion fields for correction of non-rigid motion in PET. *Annals of Nuclear Medicine*, 24(10):745–750, 2010.

[144] N. Dikaios, D. Izquierdo-Garcia, M. J. Graves, V. Mani, Z. A. Fayad, and T. D. Fryer. MRI-based motion correction of thoracic PET: initial comparison

of acquisition protocols and correction strategies suitable for simultaneous PET/MRI systems. *European Radiology*, 22(2):439–446, 2012.

[145] I. Polycarpou, C. Tsoumpas, and P. Marsden. Analysis and comparison of two methods for motion correction in PET imaging. *Medical Physics*, 39(10):6474–6483, 2012.

[146] M. D. Walker, M. Asselin, P. J. Julyan, M. Feldmann, P. Talbot, T. Jones, and J. Matthews. Bias in iterative reconstruction of low-statistics PET data: benefits of a resolution model. *Physics in Medicine and Biology*, 56(4):931, 2011.

[147] S. Y. Chun and J. A. Fessler. Noise properties of motion-compensated tomographic image reconstruction methods. *Medical Imaging, IEEE Transactions on*, 32(2):141–152, 2013.

[148] C. Tsoumpas, I. Polycarpou, K. Thielemans, C. Buerger, A. King, T. Schaeffter, and P. Marsden. The effect of regularization in motion compensated PET image reconstruction: a realistic numerical 4D simulation study. *Physics in Medicine and Biology*, 58(6):1759, 2013.

[149] A. Rahmim, P. Bloomfield, S. Houle, M. Lenox, C. Michel, K. R. Buckley, T. J. Ruth, and V. Sossi. Motion compensation in histogram-mode and list-mode EM reconstructions: beyond the event-driven approach. *Nuclear Science, IEEE Transactions on*, 51(5):2588–2596, 2004.

[150] F. Qiao, J. W. Clark Jr, T. Pan, and O. Mawlawi. Joint model of motion and anatomy for PET image reconstruction. *Medical Physics*, 34(12):4626–4639, 2007.

[151] E. Asma, R. Manjeshwar, and K. Thielemans. Theoretical comparison of motion correction techniques for PET image reconstruction. In *Nuclear Science Symposium Conference Record, 2006. IEEE*, volume 3, pages 1762–1767. IEEE, 2006.

[152] H. Fayad, H. Schmidt, C. Wuerslin, and D. Visvikis. Reconstruction-incorporated respiratory motion correction in clinical simultaneous PET/MR

imaging for oncology applications. *Journal of Nuclear Medicine*, 56(6):884–889, 2015.

[153] B. A. Mair, D. R. Gilland, and Z. Cao. Simultaneous motion estimation and image reconstruction from gated data. In *Biomedical Imaging, 2002. Proceedings. 2002 IEEE International Symposium on*, pages 661–664. IEEE, 2002.

[154] J. Qi and R. H. Huesman. List mode reconstruction for PET with motion compensation: a simulation study. In *Biomedical Imaging, 2002. Proceedings. 2002 IEEE International Symposium on*, pages 413–416. IEEE, 2002.

[155] T. Feng and B. M. Tsui. Non-rigid respiratory motion correction for 4D gated PET sinogram data. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2013 IEEE*, pages 1–5. IEEE, 2013.

[156] M. Menke, M. S. Atkins, and K. R. Buckley. Compensation methods for head motion detected during PET imaging. *Nuclear Science, IEEE Transactions on*, 43(1):310–317, 1996.

[157] L. Fin, P. Bailly, J. Daouk, and M.-E. Meyer. Motion correction based on an appropriate system matrix for statistical reconstruction of respiratory-correlated PET acquisitions. *Computer Methods and Programs in Biomedicine*, 96(3):e1–e9, 2009.

[158] P. Bühler, U. Just, E. Will, J. Kotzerke, and J. van den Hoff. An accurate method for correction of head movement in PET. *IEEE Transactions on Medical Imaging*, 23(9):1176–1185, 2004.

[159] V. W. Zhou, A. Z. Kyme, S. R. Meikle, and R. Fulton. A scheme for PET data normalization in event-based motion correction. *Physics in Medicine and Biology*, 54(17):5321, 2009.

[160] P. Manescu, H. Ladjal, J. Azencot, M. Beuve, and B. Shariat. Motion compensation for PET image reconstruction using deformable tetrahedral meshes. *Physics in Medicine and Biology*, 60(24):9269, 2015.

[161] E. Ü. Mumcuoglu, R. M. Leahy, S. R. Cherry, and E. Hoffman. Accurate geometric and physical response modelling for statistical image reconstruction in high resolution PET. In *Nuclear Science Symposium, 1996. Conference Record., 1996 IEEE*, volume 3, pages 1569–1573. IEEE, 1996.

[162] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou, and T. H. Farquhar. High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner. *Physics in Medicine and Biology*, 43(4):1001, 1998.

[163] A. J. Reader and J. Verhaeghe. 4D image reconstruction for emission tomography. *Physics in Medicine and Biology*, 59(22):R371, 2014.

[164] S. Jan, G. Santin, D. Strul, S. Staelens, K. Assie, D. Autret, S. Avner, R. Barbier, M. Bardies, P. Bloomfield, et al. GATE: a simulation toolkit for PET and SPECT. *Physics in Medicine and Biology*, 49(19):4543, 2004.

[165] W. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. Tsui. 4D XCAT phantom for multimodality imaging research. *Medical Physics*, 37(9):4902–4915, 2010.

[166] M. Dawood, F. Gigengack, X. Jiang, and K. P. Schäfers. A mass conservation-based optical flow method for cardiac motion correction in 3D-PET. *Medical physics*, 40(1):012505, 2013.

[167] P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128, 2006.

[168] J. Kybic and M. Unser. Fast parametric elastic image registration. *Image Processing, IEEE Transactions on*, 12(11):1427–1442, 2003.

[169] W. Bai and S. M. Brady. Motion correction and attenuation correction for respiratory gated PET images. *Medical Imaging, IEEE Transactions on*, 30(2):351–365, 2011.

[170] D. L. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes. Medical image registration. *Physics in Medicine and Biology*, 46(3):R1, 2001.

[171] H. J. Fayad, F. Lamare, C. C. Le Rest, V. Bettinardi, and D. Visvikis. Generation of 4-dimensional CT images based on 4-dimensional PET–derived motion fields. *Journal of Nuclear Medicine*, 54(4):631–638, 2013.

[172] F. Büther, I. Ernst, M. Dawood, P. Kraxner, M. Schäfers, O. Schober, and K. P. Schäfers. Detection of respiratory tumour motion using intrinsic list mode-driven gating in positron emission tomography. *European Journal of Nuclear Medicine and Molecular Imaging*, 37(12):2315–2327, 2010.

[173] C. Buerger, T. Schaeffter, and A. P. King. Hierarchical adaptive local affine registration for fast and robust respiratory motion estimation. *Medical Image Analysis*, 15(4):551–564, 2011.

[174] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge university press, 2012.

[175] F. Savill, T. Schaeffter, and A. P. King. Assessment of input signal positioning for cardiac respiratory motion models during different breathing patterns. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium On*, pages 1698–1701. IEEE, 2011.

[176] C. Buerger. Hierarchical adaptive local affine registration (lreg): Usage information. https://www.isd.kcl.ac.uk/internal/hyperimage/, August 2011. [Online; accessed 05-December-2016].

[177] D. Peressutti, G. P. Penney, R. J. Housden, C. Kolbitsch, A. Gomez, E.-J. Rijkhorst, D. C. Barratt, K. S. Rhode, and A. P. King. A novel Bayesian respiratory motion model to estimate and resolve uncertainty in image-guided cardiac interventions. *Medical Image Analysis*, 17(4):488–502, 2013.

[178] J. Ouyang, Q. Li, and G. El Fakhri. Magnetic resonance-based motion correction for positron emission tomography imaging. *Seminars in Nuclear Medicine*, 43(1):60 – 67, 2013. PET/MRI.

[179] M. Blume, N. Navab, and M. Rafecas. Joint image and motion reconstruction for PET using a B-spline motion model. *Physics in Medicine and Biology*, 57(24):8249, 2012.

[180] H. Ue, H. Haneishi, H. Iwanaga, and K. Suga. Nonlinear motion correction of respiratory-gated lung SPECT images. *IEEE Transactions on Medical Imaging*, 25(4):486–495, April 2006.

[181] A. Saichev and W. Woyczynski. *Distributions in the Physical and Engineering Sciences: Distributional and Fractal Calculus, Integral Transforms and Wavelets*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, 1996.

[182] D. Rees. *Essential Statistics*. Chapman and Hall, 2nd edition, 1989.

[183] D. Bailey, G. Antoch, P. Bartenstein, H. Barthel, A. Beer, S. Bisdas, D. Bluemke, R. Boellaard, C. Claussen, C. Franzius, et al. Combined PET/MR: the real work has just started. summary report of the third international workshop on PET/MR imaging; February 17–21, 2014, Tübingen, Germany. *Molecular Imaging and Biology*, 17(3):297–312, 2015.

[184] T. Pan, O. Mawlawi, S. A. Nehmeh, Y. E. Erdi, D. Luo, H. H. Liu, R. Castillo, R. Mohan, Z. Liao, and H. Macapinlac. Attenuation correction of PET images with respiration-averaged CT images in PET/CT. *Journal of Nuclear Medicine*, 46(9):1481–1487, 2005.

[185] T. M. Blodgett, A. S. Mehta, A. S. Mehta, C. M. Laymon, J. Carney, and D. W. Townsend. PET/CT artifacts. *Clinical Imaging*, 35(1):49–63, 2011.

[186] A. Rahmim, J. Tang, and H. Zaidi. Four-dimensional image reconstruction strategies in cardiac-gated and respiratory-gated PET imaging. *PET Clinics*, 8(1):51–67, 2013.